

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
4 January 2001 (04.01.2001)

PCT

(10) International Publication Number  
**WO 01/00842 A2**

(51) International Patent Classification<sup>7</sup>: C12N 15/31, 15/52, 9/00, C07K 14/34 199 42 088.2 3 September 1999 (03.09.1999) DE

(21) International Application Number: PCT/IB00/00911

(71) Applicant: BASF AKTIENGESELLSCHAFT  
[DE/DE]; D-67056 Ludwigshafen (DE).

(22) International Filing Date: 23 June 2000 (23.06.2000)

(72) Inventors: POMPEJUS, Markus; Wenjenstrasse 21, D-67251 Freinsheim (DE). KRÖGER, Burkhard; Im Waldhof 1, D-67117 Limburgerhof (DE). SCHRÖDER, Hartwig; Goethestrasse 5, D-69226 Nussloch (DE). ZELDER, Oskar; Rossmarktstrasse 27, D-67346 Speyer (DE). HABERHAUER, Gregor; Moselstrasse 42, D-67117 Limburgerhof (DE).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/141,031	25 June 1999 (25.06.1999)	US
199 31 636.8	8 July 1999 (08.07.1999)	DE
199 32 125.6	9 July 1999 (09.07.1999)	DE
199 32 126.4	9 July 1999 (09.07.1999)	DE
199 32 127.2	9 July 1999 (09.07.1999)	DE
199 32 128.0	9 July 1999 (09.07.1999)	DE
199 32 129.9	9 July 1999 (09.07.1999)	DE
199 32 226.0	9 July 1999 (09.07.1999)	DE
199 32 920.6	14 July 1999 (14.07.1999)	DE
199 32 922.2	14 July 1999 (14.07.1999)	DE
199 32 924.9	14 July 1999 (14.07.1999)	DE
199 32 928.1	14 July 1999 (14.07.1999)	DE
199 32 930.3	14 July 1999 (14.07.1999)	DE
199 32 933.8	14 July 1999 (14.07.1999)	DE
199 32 935.4	14 July 1999 (14.07.1999)	DE
199 32 973.7	14 July 1999 (14.07.1999)	DE
199 33 002.6	14 July 1999 (14.07.1999)	DE
199 33 003.4	14 July 1999 (14.07.1999)	DE
199 33 005.0	14 July 1999 (14.07.1999)	DE
199 33 006.9	14 July 1999 (14.07.1999)	DE
199 41 378.9	31 August 1999 (31.08.1999)	DE
199 41 379.7	31 August 1999 (31.08.1999)	DE
199 41 390.8	31 August 1999 (31.08.1999)	DE
199 41 391.6	31 August 1999 (31.08.1999)	DE

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— Without international search report and to be republished upon receipt of that report.

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS INVOLVED IN HOMEOSTASIS AND ADAPTATION

(57) Abstract: Isolated nucleic acid molecules, designated HA nucleic acid molecules, which encode novel HA proteins from *Corynebacterium glutamicum* are described. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing HA nucleic acid molecules, and host cells into which the expression vectors have been introduced. The invention still further provides isolated HA proteins, mutated HA proteins, fusion proteins, antigenic peptides and methods for the improvement of production of a desired compound from *C. glutamicum* based on genetic engineering of HA genes in this organism.

WO 01/00842 A2

***CORYNEBACTERIUM GLUTAMICUM* GENES ENCODING PROTEINS  
INVOLVED IN HOMEOSTASIS AND ADAPTATION**

**Related Applications**

5           This application claims priority to prior filed U.S. Provisional Patent Application  
Serial No. 60/141031, filed June 25, 1999. This application also claims priority to prior  
filed German Patent Application No. 19931636.8, filed July 8, 1999, German Patent  
Application No. 19932125.6, filed July 9, 1999, German Patent Application No.  
19932126.4, filed July 9, 1999, German Patent Application No. 19932127.2, filed July  
10 9, 1999, German Patent Application No. 19932128.0, filed July 9, 1999, German Patent  
Application No. 19932129.9, filed July 9, 1999, German Patent Application No.  
19932226.0, filed July 9, 1999, German Patent Application No. 19932920.6, filed July  
14, 1999, German Patent Application No. 19932922.2, filed July 14, 1999, German  
Patent Application No. 19932924.9, filed July 14, 1999, German Patent Application No.  
15 19932928.1, filed July 14, 1999, German Patent Application No. 19932930.3, filed July  
14, 1999, German Patent Application No. 19932933.8, filed July 14, 1999, German  
Patent Application No. 19932935.4, filed July 14, 1999, German Patent Application No.  
19932973.7, filed July 14, 1999, German Patent Application No. 19933002.6, filed July  
14, 1999, German Patent Application No. 19933003.4, filed July 14, 1999, German  
20 Patent Application No. 19933005.0, filed July 14, 1999, German Patent Application No.  
19933006.9, filed July 14, 1999, German Patent Application No. 19941378.9, filed  
August 31, 1999, German Patent Application No. 19941379.7, filed August 31, 1999,  
German Patent Application No. 19941390.8, filed August 31, 1999, German Patent  
Application No. 19941391.6, filed August 31, 1999, and German Patent Application No.  
25 19942088.2, filed September 3, 1999. The entire contents of all of the aforementioned  
applications are hereby expressly incorporated herein by this reference.

**Background of the Invention**

30           Certain products and by-products of naturally-occurring metabolic processes in  
cells have utility in a wide array of industries, including the food, feed, cosmetics, and  
pharmaceutical industries. These molecules, collectively termed 'fine chemicals',  
include organic acids, both proteinogenic and non-proteinogenic amino acids,

- 2 -

nucleotides and nucleosides, lipids and fatty acids, diols, carbohydrates, aromatic compounds, vitamins and cofactors, and enzymes. Their production is most conveniently performed through the large-scale culture of bacteria developed to produce and secrete large quantities of one or more desired molecules. One particularly useful  
5 organism for this purpose is *Corynebacterium glutamicum*, a gram positive, nonpathogenic bacterium. Through strain selection, a number of mutant strains have been developed which produce an array of desirable compounds. However, selection of strains improved for the production of a particular molecule is a time-consuming and difficult process.

10

### Summary of the Invention

The invention provides novel bacterial nucleic acid molecules which have a variety of uses. These uses include the identification of microorganisms which can be used to produce fine chemicals, the modulation of fine chemical production in *C.*  
15 *glutamicum* or related bacteria, the typing or identification of *C. glutamicum* or related bacteria, as reference points for mapping the *C. glutamicum* genome, and as markers for transformation. These novel nucleic acid molecules encode proteins, referred to herein as homeostasis and adaptation (HA) proteins.

*C. glutamicum* is a gram positive, aerobic bacterium which is commonly used in  
20 industry for the large-scale production of a variety of fine chemicals, and also for the degradation of hydrocarbons (such as in petroleum spills) and for the oxidation of terpenoids. The HA nucleic acid molecules of the invention, therefore, can be used to identify microorganisms which can be used to produce fine chemicals, e.g., by fermentation processes. Modulation of the expression of the HA nucleic acids of the  
25 invention, or modification of the sequence of the HA nucleic acid molecules of the invention, can be used to modulate the production of one or more fine chemicals from a microorganism (e.g., to improve the yield or production of one or more fine chemicals from a *Corynebacterium* or *Brevibacterium* species).

The HA nucleic acids of the invention may also be used to identify an organism  
30 as being *Corynebacterium glutamicum* or a close relative thereof, or to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C.*

*glutamicum* genes; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is nonpathogenic, it is related to species pathogenic in humans, such as *Corynebacterium diphtheriae* (the causative agent of diphtheria); the detection of such organisms is of significant clinical relevance.

The HA nucleic acid molecules of the invention may also serve as reference points for mapping of the *C. glutamicum* genome, or of genomes of related organisms. Similarly, these molecules, or variants or portions thereof, may serve as markers for genetically engineered *Corynebacterium* or *Brevibacterium* species.

The HA proteins encoded by the novel nucleic acid molecules of the invention are capable of, for example, performing a function involved in the maintenance of homeostasis in *C. glutamicum*, or in the ability of this microorganism to adapt to different environmental conditions. Given the availability of cloning vectors for use in *Corynebacterium glutamicum*, such as those disclosed in Sinskey *et al.*, U.S. Patent No. 4,649,119, and techniques for genetic manipulation of *C. glutamicum* and the related *Brevibacterium* species (*e.g.*, *lactofermentum*) (Yoshihama *et al.*, *J. Bacteriol.* 162: 591-597 (1985); Katsumata *et al.*, *J. Bacteriol.* 159: 306-311 (1984); and Santamaria *et al.*, *J. Gen. Microbiol.* 130: 2237-2246 (1984)), the nucleic acid molecules of the invention may be utilized in the genetic engineering of this organism to make it a better or more efficient producer of one or more fine chemicals. This improved production or efficiency of production of a fine chemical may be due to a direct effect of manipulation of a gene of the invention, or it may be due to an indirect effect of such manipulation.

There are a number of mechanisms by which the alteration of an HA protein of the invention may directly affect the yield, production, and/or efficiency of production of a fine chemical from a *C. glutamicum* strain incorporating such an altered protein. For example, by engineering enzymes which modify or degrade aromatic or aliphatic compounds such that these enzymes are increased or decreased in activity or number, it may be possible to modulate the production of one or more fine chemicals which are the modification or degradation products of these compounds. Similarly, enzymes involved in the metabolism of inorganic compounds provide key molecules (*e.g.* phosphorous,



- 4 -

sulfur, and nitrogen molecules) for the biosynthesis of such fine chemicals as amino acids, vitamins, and nucleic acids. By altering the activity or number of these enzymes in *C. glutamicum*, it may be possible to increase the conversion of these inorganic compounds (or to use alternate inorganic compounds) to thus permit improved rates of incorporation of inorganic atoms into these fine chemicals. Genetic engineering of *C. glutamicum* enzymes involved in general cellular processes may also directly improve fine chemical production, since many of these enzymes directly modify fine chemicals (e.g., amino acids) or the enzymes which are involved in fine chemical synthesis or secretion. Modulation of the activity or number of cellular proteases may also have a direct effect on fine chemical production, since many proteases may degrade fine chemicals or enzymes involved in fine chemical production or breakdown.

Further, the aforementioned enzymes which participate in aromatic/aliphatic compound modification or degradation, general biocatalysis, inorganic compound metabolism or proteolysis are each themselves fine chemicals, desirable for their activity in various *in vitro* industrial applications. By altering the number of copies of the gene for one or more of these enzymes in *C. glutamicum* it may be possible to increase the number of these proteins produced by the cell, thereby increasing the potential yield or efficiency of production of these proteins from large-scale *C. glutamicum* or related bacterial cultures.

The alteration of an HA protein of the invention may also indirectly affect the yield, production, and/or efficiency of production of a fine chemical from a *C. glutamicum* strain incorporating such an altered protein. For example, by modulating the activity and/or number of those proteins involved in the construction or rearrangement of the cell wall, it may be possible to modify the structure of the cell wall itself such that the cell is able to better withstand the mechanical and other stresses present during large-scale fermentative culture. Also, large-scale growth of *C. glutamicum* requires significant cell wall production. Modulation of the activity or number of cell wall biosynthetic or degradative enzymes may allow more rapid rates of cell wall biosynthesis, which in turn may permit increased growth rates of this microorganism in culture and thereby increase the number of cells producing the desired fine chemical.

- 5 -

By modifying the HA enzymes of the invention, one may also indirectly impact the yield, production, or efficiency of production of one or more fine chemicals from *C. glutamicum*. For example, many of the general enzymes in *C. glutamicum* may have a significant impact on global cellular processes (e.g., regulatory processes) which in turn have a significant effect on fine chemical metabolism. Similarly, proteases, enzymes which modify or degrade possibly toxic aromatic or aliphatic compounds, and enzymes which promote the metabolism of inorganic compounds all serve to increase the viability of *C. glutamicum*. The proteases aid in the selective removal of misfolded or misregulated proteins, such as those that might occur under the relatively stressful environmental conditions encountered during large-scale fermentor culture. By altering these proteins, it may be possible to further enhance this activity and to improve the viability of *C. glutamicum* in culture. The aromatic/aliphatic modification or degradation proteins not only serve to detoxify these waste compounds (which may be encountered as impurities in culture medium or as waste products from cells themselves), but also to permit the cells to utilize alternate carbon sources if the optimal carbon source is limiting in the culture. By increasing their number and/or activity, the survival of *C. glutamicum* cells in culture may be enhanced. The inorganic metabolism proteins of the invention supply the cell with inorganic molecules required for all protein and nucleotide (among others) synthesis, and thus are critical for the overall viability of the cell. An increase in the number of viable cells producing one or more desired fine chemicals in large-scale culture should result in a concomitant increase in the yield, production, and/or efficiency of production of the fine chemical in the culture.

The invention provides novel nucleic acid molecules which encode proteins, referred to herein as HA proteins, which are capable of, for example, performing a function involved in the maintenance of homeostasis in *C. glutamicum*, or of participating in the ability of this microorganism to adapt to different environmental conditions. Nucleic acid molecules encoding an HA protein are referred to herein as HA nucleic acid molecules. In a preferred embodiment, an HA protein participates in *C. glutamicum* cell wall biosynthesis or rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or possesses a *C. glutamicum* enzymatic or proteolytic activity. Examples of such proteins include those encoded by the genes set forth in Table 1.

- 6 -

Accordingly, one aspect of the invention pertains to isolated nucleic acid molecules (*e.g.*, cDNAs, DNAs, or RNAs) comprising a nucleotide sequence encoding an HA protein or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection or amplification of HA-  
5 encoding nucleic acids (*e.g.*, DNA or mRNA). In particularly preferred embodiments, the isolated nucleic acid molecule comprises one of the nucleotide sequences set forth as the odd-numbered SEQ ID NOs in the Sequence Listing (*e.g.*, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7...), or the coding region or a complement thereof of one of these nucleotide sequences. In other particularly preferred embodiments, the  
10 isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes to or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80% or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence set forth as an odd-numbered SEQ ID NO in the Sequence Listing (*e.g.*, SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5,  
15 SEQ ID NO:7...), or a portion thereof. In other preferred embodiments, the isolated nucleic acid molecule encodes one of the amino acid sequences set forth in as an even-numbered SEQ ID NO in the Sequence Listing (*e.g.*, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8...). The preferred HA proteins of the present invention also preferably possess at least one of the HA activities described herein.

20 In another embodiment, the isolated nucleic acid molecule encodes a protein or portion thereof wherein the protein or portion thereof includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence having an even-numbered SEQ ID NO: in the Sequence Listing), *e.g.*, sufficiently homologous to an amino acid sequence of the invention such that the protein  
25 or portion thereof maintains an HA activity. Preferably, the protein or portion thereof encoded by the nucleic acid molecule maintains the ability to participate in the maintenance of homeostasis in *C. glutamicum*, or to perform a function involved in the adaptation of this microorganism to different environmental conditions. In one embodiment, the protein encoded by the nucleic acid molecule is at least about 50%,  
30 preferably at least about 60%, and more preferably at least about 70%, 80%, or 90% and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an amino acid sequence of the invention (*e.g.*, an entire amino acid sequence selected from

- 7 -

those having an even-numbered SEQ ID NO in the Sequence Listing). In another preferred embodiment, the protein is a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of the invention (encoded by an open reading frame shown in the corresponding odd-numbered SEQ ID NOs in the Sequence Listing (e.g., SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7....)).

In another preferred embodiment, the isolated nucleic acid molecule is derived from *C. glutamicum* and encodes a protein (e.g., an HA fusion protein) which includes a biologically active domain which is at least about 50% or more homologous to one of the amino acid sequences of the invention (e.g., a sequence of one of the even-numbered SEQ ID NOs in the Sequence Listing) and is able to participate in the repair or recombination of DNA, in the transposition of genetic material, in gene expression (i.e., the processes of transcription or translation), in protein folding, or in protein secretion in *Corynebacterium glutamicum*, or has one or more of the activities set forth in Table 1, and which also includes heterologous nucleic acid sequences encoding a heterologous polypeptide or regulatory regions.

In another embodiment, the isolated nucleic acid molecule is at least 15 nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising a nucleotide sequence of the invention (e.g., a sequence of an odd-numbered SEQ ID NO in the Sequence Listing). Preferably, the isolated nucleic acid molecule corresponds to a naturally-occurring nucleic acid molecule. More preferably, the isolated nucleic acid encodes a naturally-occurring *C. glutamicum* HA protein, or a biologically active portion thereof.

Another aspect of the invention pertains to vectors, e.g., recombinant expression vectors, containing the nucleic acid molecules of the invention, and host cells into which such vectors have been introduced. In one embodiment, such a host cell is used to produce an HA protein by culturing the host cell in a suitable medium. The HA protein can be then isolated from the medium or the host cell.

Yet another aspect of the invention pertains to a genetically altered microorganism in which an HA gene has been introduced or altered. In one embodiment, the genome of the microorganism has been altered by introduction of a nucleic acid molecule of the invention encoding wild-type or mutated HA sequence as a transgene. In another embodiment, an endogenous HA gene within the genome of the

- 8 -

microorganism has been altered, *e.g.*, functionally disrupted, by homologous recombination with an altered HA gene. In another embodiment, an endogenous or introduced HA gene in a microorganism has been altered by one or more point mutations, deletions, or inversions, but still encodes a functional HA protein. In still  
5 another embodiment, one or more of the regulatory regions (*e.g.*, a promoter, repressor, or inducer) of an HA gene in a microorganism has been altered (*e.g.*, by deletion, truncation, inversion, or point mutation) such that the expression of the HA gene is modulated. In a preferred embodiment, the microorganism belongs to the genus *Corynebacterium* or *Brevibacterium*, with *Corynebacterium glutamicum* being  
10 particularly preferred. In a preferred embodiment, the microorganism is also utilized for the production of a desired compound, such as an amino acid, with lysine being particularly preferred.

In another aspect, the invention provides a method of identifying the presence or activity of *Corynebacterium diphtheriae* in a subject. This method includes detection of  
15 one or more of the nucleic acid or amino acid sequences of the invention (*e.g.*, the sequences set forth in the Sequence Listing as SEQ ID NOs 1 through 440) in a subject, thereby detecting the presence or activity of *Corynebacterium diphtheriae* in the subject.

Still another aspect of the invention pertains to an isolated HA protein or a portion, *e.g.*, a biologically active portion, thereof. In a preferred embodiment, the  
20 isolated HA protein or portion thereof can participate in the maintenance of homeostasis in *C. glutamicum*, or can perform a function involved in the adaptation of this microorganism to different environmental conditions. In another preferred embodiment, the isolated HA protein or portion thereof is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: in the  
25 Sequence Listing) such that the protein or portion thereof maintains the ability to participate in the maintenance of homeostasis in *C. glutamicum*, or to perform a function involved in the adaptation of this microorganism to different environmental conditions.

The invention also provides an isolated preparation of an HA protein. In preferred embodiments, the HA protein comprises an amino acid sequence of the  
30 invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing). In another preferred embodiment, the invention pertains to an isolated full length protein which is substantially homologous to an entire amino acid sequence of the invention

- 9 -

(*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) (encoded by an open reading frame set forth in a corresponding odd-numbered SEQ ID NO: of the Sequence Listing). In yet another embodiment, the protein is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90%,  
5 and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an entire amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing). In other embodiments, the isolated HA protein comprises an amino acid sequence which is at least about 50% or more homologous to one of the amino acid sequences of the invention (*e.g.*, a sequence of an even-numbered  
10 SEQ ID NO: of the Sequence Listing) and is able to participate in the maintenance of homeostasis in *C. glutamicum*, or to perform a function involved in the adaptation of this microorganism to different environmental conditions, or has one or more of the activities set forth in Table 1.

Alternatively, the isolated HA protein can comprise an amino acid sequence  
15 which is encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80%, or 90%, and even more preferably at least about 95%, 96%, 97%, 98,%, or 99% or more homologous, to a nucleotide sequence of one of the even-numbered SEQ ID NOs set forth in the Sequence Listing. It is also preferred  
20 that the preferred forms of HA proteins also have one or more of the HA bioactivities described herein.

The HA polypeptide, or a biologically active portion thereof, can be operatively linked to a non-HA polypeptide to form a fusion protein. In preferred embodiments, this fusion protein has an activity which differs from that of the HA protein alone. In other  
25 preferred embodiments, this fusion protein participates in the maintenance of homeostasis in *C. glutamicum*, or performs a function involved in the adaptation of this microorganism to different environmental conditions. In particularly preferred embodiments, integration of this fusion protein into a host cell modulates production of a desired compound from the cell.

30 In another aspect, the invention provides methods for screening molecules which modulate the activity of an HA protein, either by interacting with the protein itself or a

- 10 -

substrate or binding partner of the HA protein, or by modulating the transcription or translation of an HA nucleic acid molecule of the invention.

Another aspect of the invention pertains to a method for producing a fine chemical. This method involves the culturing of a cell containing a vector directing the expression of an HA nucleic acid molecule of the invention, such that a fine chemical is produced. In a preferred embodiment, this method further includes the step of obtaining a cell containing such a vector, in which a cell is transfected with a vector directing the expression of an HA nucleic acid. In another preferred embodiment, this method further includes the step of recovering the fine chemical from the culture. In a particularly preferred embodiment, the cell is from the genus *Corynebacterium* or *Brevibacterium*, or is selected from those strains set forth in Table 3.

Another aspect of the invention pertains to methods for modulating production of a molecule from a microorganism. Such methods include contacting the cell with an agent which modulates HA protein activity or HA nucleic acid expression such that a cell associated activity is altered relative to this same activity in the absence of the agent. In a preferred embodiment, the cell is modulated for one or more *C. glutamicum* processes involved in cell wall biosynthesis or rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or enzymatic or proteolytic activities. The agent which modulates HA protein activity can be an agent which stimulates HA protein activity or HA nucleic acid expression. Examples of agents which stimulate HA protein activity or HA nucleic acid expression include small molecules, active HA proteins, and nucleic acids encoding HA proteins that have been introduced into the cell. Examples of agents which inhibit HA activity or expression include small molecules and antisense HA nucleic acid molecules.

Another aspect of the invention pertains to methods for modulating yields of a desired compound from a cell, involving the introduction of a wild-type or mutant HA gene into a cell, either maintained on a separate plasmid or integrated into the genome of the host cell. If integrated into the genome, such integration can be random, or it can take place by homologous recombination such that the native gene is replaced by the introduced copy, causing the production of the desired compound from the cell to be modulated. In a preferred embodiment, said yields are increased. In another preferred embodiment, said chemical is a fine chemical. In a particularly preferred embodiment,

- 11 -

said fine chemical is an amino acid. In especially preferred embodiments, said amino acid is L-lysine.

### **Detailed Description of the Invention**

5           The present invention provides HA nucleic acid and protein molecules which are involved in *C. glutamicum* cell wall biosynthesis or rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or that have a *C. glutamicum* enzymatic or proteolytic activity. The molecules of the invention may be utilized in the modulation of production of fine chemicals from  
10 microorganisms, such as *C. glutamicum*, either directly (*e.g.*, where overexpression or optimization of activity of a protein involved in the production of a fine chemical (*e.g.*, an enzyme) has a direct impact on the yield, production, and/or efficiency of production of a fine chemical from the modified *C. glutamicum*), or an indirect impact which nonetheless results in an increase of yield, production, and/or efficiency of production of  
15 the desired compound (*e.g.*, where modulation of the activity or number of copies of a *C. glutamicum* aromatic or aliphatic modification or degradation protein results in an increase in the viability of *C. glutamicum* cells, which in turn permits increased production in a large-scale culture setting). Aspects of the invention are further explicated below.

20

#### **I. Fine Chemicals**

          The term 'fine chemical' is art-recognized and includes molecules produced by an organism which have applications in various industries, such as, but not limited to, the pharmaceutical, agriculture, and cosmetics industries. Such compounds include  
25 organic acids, such as tartaric acid, itaconic acid, and diaminopimelic acid, both proteinogenic and non-proteinogenic amino acids, purine and pyrimidine bases, nucleosides, and nucleotides (as described *e.g.* in Kuninaka, A. (1996) Nucleotides and related compounds, p. 561-612, in Biotechnology vol. 6, Rehm *et al.*, eds. VCH: Weinheim, and references contained therein), lipids, both saturated and unsaturated fatty  
30 acids (*e.g.*, arachidonic acid), diols (*e.g.*, propane diol, and butane diol), carbohydrates (*e.g.*, hyaluronic acid and trehalose), aromatic compounds (*e.g.*, aromatic amines, vanillin, and indigo), vitamins and cofactors (as described in Ullmann's Encyclopedia of



Industrial Chemistry, vol. A27, "Vitamins", p. 443-613 (1996) VCH: Weinheim and references therein; and Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research – Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press, (1995)), enzymes, polyketides (Cane *et al.* (1998) *Science* 282: 63-68), and all other chemicals described in Gutcho (1983) *Chemicals by Fermentation*, Noyes Data Corporation, ISBN: 0818805086 and references therein. The metabolism and uses of certain of these fine chemicals are further explicated below.

10

#### A. *Amino Acid Metabolism and Uses*

Amino acids comprise the basic structural units of all proteins, and as such are essential for normal cellular functioning in all organisms. The term "amino acid" is art-recognized. The proteinogenic amino acids, of which there are 20 species, serve as structural units for proteins, in which they are linked by peptide bonds, while the nonproteinogenic amino acids (hundreds of which are known) are not normally found in proteins (see Ulmann's *Encyclopedia of Industrial Chemistry*, vol. A2, p. 57-97 VCH: Weinheim (1985)). Amino acids may be in the D- or L- optical configuration, though L-amino acids are generally the only type found in naturally-occurring proteins.

Biosynthetic and degradative pathways of each of the 20 proteinogenic amino acids have been well characterized in both prokaryotic and eukaryotic cells (see, for example, Stryer, L. *Biochemistry*, 3<sup>rd</sup> edition, pages 578-590 (1988)). The 'essential' amino acids (histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine), so named because they are generally a nutritional requirement due to the complexity of their biosyntheses, are readily converted by simple biosynthetic pathways to the remaining 11 'nonessential' amino acids (alanine, arginine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, serine, and tyrosine). Higher animals do retain the ability to synthesize some of these amino acids, but the essential amino acids must be supplied from the diet in order for normal protein synthesis to occur.

30

Aside from their function in protein biosynthesis, these amino acids are interesting chemicals in their own right, and many have been found to have various applications in the food, feed, chemical, cosmetics, agriculture, and pharmaceutical

industries. Lysine is an important amino acid in the nutrition not only of humans, but also of monogastric animals such as poultry and swine. Glutamate is most commonly used as a flavor additive (mono-sodium glutamate, MSG) and is widely used throughout the food industry, as are aspartate, phenylalanine, glycine, and cysteine. Glycine, L-methionine and tryptophan are all utilized in the pharmaceutical industry. Glutamine, valine, leucine, isoleucine, histidine, arginine, proline, serine and alanine are of use in both the pharmaceutical and cosmetics industries. Threonine, tryptophan, and D/L-methionine are common feed additives. (Leuchtenberger, W. (1996) Amino acids – technical production and use, p. 466-502 in Rehm *et al.* (eds.) Biotechnology vol. 6, chapter 14a, VCH: Weinheim). Additionally, these amino acids have been found to be useful as precursors for the synthesis of synthetic amino acids and proteins, such as N-acetylcysteine, S-carboxymethyl-L-cysteine, (S)-5-hydroxytryptophan, and others described in Ulmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97, VCH: Weinheim, 1985.

The biosynthesis of these natural amino acids in organisms capable of producing them, such as bacteria, has been well characterized (for review of bacterial amino acid biosynthesis and regulation thereof, see Umbarger, H.E.(1978) *Ann. Rev. Biochem.* 47: 533-606). Glutamate is synthesized by the reductive amination of  $\alpha$ -ketoglutarate, an intermediate in the citric acid cycle. Glutamine, proline, and arginine are each subsequently produced from glutamate. The biosynthesis of serine is a three-step process beginning with 3-phosphoglycerate (an intermediate in glycolysis), and resulting in this amino acid after oxidation, transamination, and hydrolysis steps. Both cysteine and glycine are produced from serine; the former by the condensation of homocysteine with serine, and the latter by the transferal of the side-chain  $\beta$ -carbon atom to tetrahydrofolate, in a reaction catalyzed by serine transhydroxymethylase. Phenylalanine, and tyrosine are synthesized from the glycolytic and pentose phosphate pathway precursors erythrose 4-phosphate and phosphoenolpyruvate in a 9-step biosynthetic pathway that differ only at the final two steps after synthesis of prephenate. Tryptophan is also produced from these two initial molecules, but its synthesis is an 11-step pathway. Tyrosine may also be synthesized from phenylalanine, in a reaction catalyzed by phenylalanine hydroxylase. Alanine, valine, and leucine are all biosynthetic products of pyruvate, the final product of glycolysis. Aspartate is formed

- 14 -

from oxaloacetate, an intermediate of the citric acid cycle. Asparagine, methionine, threonine, and lysine are each produced by the conversion of aspartate. Isoleucine is formed from threonine. A complex 9-step pathway results in the production of histidine from 5-phosphoribosyl-1-pyrophosphate, an activated sugar.

5       Amino acids in excess of the protein synthesis needs of the cell cannot be stored, and are instead degraded to provide intermediates for the major metabolic pathways of the cell (for review see Stryer, L. Biochemistry 3<sup>rd</sup> ed. Ch. 21 "Amino Acid Degradation and the Urea Cycle" p. 495-516 (1988)). Although the cell is able to convert unwanted amino acids into useful metabolic intermediates, amino acid production is costly in  
10   terms of energy, precursor molecules, and the enzymes necessary to synthesize them. Thus it is not surprising that amino acid biosynthesis is regulated by feedback inhibition, in which the presence of a particular amino acid serves to slow or entirely stop its own production (for overview of feedback mechanisms in amino acid biosynthetic pathways, see Stryer, L. Biochemistry, 3<sup>rd</sup> ed. Ch. 24: "Biosynthesis of Amino Acids and Heme" p.  
15   575-600 (1988)). Thus, the output of any particular amino acid is limited by the amount of that amino acid present in the cell.

#### *B. Vitamin, Cofactor, and Nutraceutical Metabolism and Uses*

Vitamins, cofactors, and nutraceuticals comprise another group of molecules  
20   which the higher animals have lost the ability to synthesize and so must ingest, although they are readily synthesized by other organisms such as bacteria. These molecules are either bioactive substances themselves, or are precursors of biologically active substances which may serve as electron carriers or intermediates in a variety of metabolic pathways. Aside from their nutritive value, these compounds also have  
25   significant industrial value as coloring agents, antioxidants, and catalysts or other processing aids. (For an overview of the structure, activity, and industrial applications of these compounds, see, for example, Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996.) The term "vitamin" is art-recognized, and includes nutrients which are required by an organism for normal  
30   functioning, but which that organism cannot synthesize by itself. The group of vitamins may encompass cofactors and nutraceutical compounds. The language "cofactor" includes nonproteinaceous compounds required for a normal enzymatic activity to

occur. Such compounds may be organic or inorganic; the cofactor molecules of the invention are preferably organic. The term "nutraceutical" includes dietary supplements having health benefits in plants and animals, particularly humans. Examples of such molecules are vitamins, antioxidants, and also certain lipids (*e.g.*, polyunsaturated fatty acids).

The biosynthesis of these molecules in organisms capable of producing them, such as bacteria, has been largely characterized (Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613, VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research – Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press: Champaign, IL X, 374 S).

Thiamin (vitamin B<sub>1</sub>) is produced by the chemical coupling of pyrimidine and thiazole moieties. Riboflavin (vitamin B<sub>2</sub>) is synthesized from guanosine-5'-triphosphate (GTP) and ribose-5'-phosphate. Riboflavin, in turn, is utilized for the synthesis of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). The family of compounds collectively termed 'vitamin B<sub>6</sub>' (*e.g.*, pyridoxine, pyridoxamine, pyridoxal-5'-phosphate, and the commercially used pyridoxin hydrochloride) are all derivatives of the common structural unit, 5-hydroxy-6-methylpyridine. Pantothenate (pantothenic acid, (R)-(+)-N-(2,4-dihydroxy-3,3-dimethyl-1-oxobutyl)- $\beta$ -alanine) can be produced either by chemical synthesis or by fermentation. The final steps in pantothenate biosynthesis consist of the ATP-driven condensation of  $\beta$ -alanine and pantoic acid. The enzymes responsible for the biosynthesis steps for the conversion to pantoic acid, to  $\beta$ -alanine and for the condensation to pantothenic acid are known. The metabolically active form of pantothenate is Coenzyme A, for which the biosynthesis proceeds in 5 enzymatic steps. Pantothenate, pyridoxal-5'-phosphate, cysteine and ATP are the precursors of Coenzyme A. These enzymes not only catalyze the formation of pantothenate, but also the production of (R)-pantoic acid, (R)-pantolacton, (R)-panthanol (provitamin B<sub>5</sub>), pantetheine (and its derivatives) and coenzyme A.

Biotin biosynthesis from the precursor molecule pimeloyl-CoA in microorganisms has been studied in detail and several of the genes involved have been

identified. Many of the corresponding proteins have been found to also be involved in Fe-cluster synthesis and are members of the *nifS* class of proteins. Lipoic acid is derived from octanoic acid, and serves as a coenzyme in energy metabolism, where it becomes part of the pyruvate dehydrogenase complex and the  $\alpha$ -ketoglutarate

5 dehydrogenase complex. The folates are a group of substances which are all derivatives of folic acid, which in turn is derived from L-glutamic acid, p-amino-benzoic acid and 6-methylpterin. The biosynthesis of folic acid and its derivatives, starting from the metabolism intermediates guanosine-5'-triphosphate (GTP), L-glutamic acid and p-amino-benzoic acid has been studied in detail in certain microorganisms.

10       Corrinoids (such as the cobalamines and particularly vitamin B<sub>12</sub>) and porphyrines belong to a group of chemicals characterized by a tetrapyrrole ring system. The biosynthesis of vitamin B<sub>12</sub> is sufficiently complex that it has not yet been completely characterized, but many of the enzymes and substrates involved are now known. Nicotinic acid (nicotinate), and nicotinamide are pyridine derivatives which are  
15 also termed 'niacin'. Niacin is the precursor of the important coenzymes NAD (nicotinamide adenine dinucleotide) and NADP (nicotinamide adenine dinucleotide phosphate) and their reduced forms.

      The large-scale production of these compounds has largely relied on cell-free chemical syntheses, though some of these chemicals have also been produced by large-  
20 scale culture of microorganisms, such as riboflavin, Vitamin B<sub>6</sub>, pantothenate, and biotin. Only Vitamin B<sub>12</sub> is produced solely by fermentation, due to the complexity of its synthesis. *In vitro* methodologies require significant inputs of materials and time, often at great cost.

### 25   C. Purine, Pyrimidine, Nucleoside and Nucleotide Metabolism and Uses

      Purine and pyrimidine metabolism genes and their corresponding proteins are important targets for the therapy of tumor diseases and viral infections. The language "purine" or "pyrimidine" includes the nitrogenous bases which are constituents of nucleic acids, co-enzymes, and nucleotides. The term "nucleotide" includes the basic  
30 structural units of nucleic acid molecules, which are comprised of a nitrogenous base, a pentose sugar (in the case of RNA, the sugar is ribose; in the case of DNA, the sugar is D-deoxyribose), and phosphoric acid. The language "nucleoside" includes molecules

which serve as precursors to nucleotides, but which are lacking the phosphoric acid moiety that nucleotides possess. By inhibiting the biosynthesis of these molecules, or their mobilization to form nucleic acid molecules, it is possible to inhibit RNA and DNA synthesis; by inhibiting this activity in a fashion targeted to cancerous cells, the ability of tumor cells to divide and replicate may be inhibited. Additionally, there are nucleotides which do not form nucleic acid molecules, but rather serve as energy stores (*i.e.*, AMP) or as coenzymes (*i.e.*, FAD and NAD).

Several publications have described the use of these chemicals for these medical indications, by influencing purine and/or pyrimidine metabolism (*e.g.* Christopherson, R.I. and Lyons, S.D. (1990) "Potent inhibitors of *de novo* pyrimidine and purine biosynthesis as chemotherapeutic agents." *Med. Res. Reviews* 10: 505-548). Studies of enzymes involved in purine and pyrimidine metabolism have been focused on the development of new drugs which can be used, for example, as immunosuppressants or anti-proliferants (Smith, J.L., (1995) "Enzymes in nucleotide synthesis." *Curr. Opin. Struct. Biol.* 5: 752-757; (1995) *Biochem Soc. Transact.* 23: 877-902). However, purine and pyrimidine bases, nucleosides and nucleotides have other utilities: as intermediates in the biosynthesis of several fine chemicals (*e.g.*, thiamine, S-adenosyl-methionine, folates, or riboflavin), as energy carriers for the cell (*e.g.*, ATP or GTP), and for chemicals themselves, commonly used as flavor enhancers (*e.g.*, IMP or GMP) or for several medicinal applications (see, for example, Kuninaka, A. (1996) Nucleotides and Related Compounds in Biotechnology vol. 6, Rehm *et al.*, eds. VCH: Weinheim, p. 561-612). Also, enzymes involved in purine, pyrimidine, nucleoside, or nucleotide metabolism are increasingly serving as targets against which chemicals for crop protection, including fungicides, herbicides and insecticides, are developed.

The metabolism of these compounds in bacteria has been characterized (for reviews see, for example, Zalkin, H. and Dixon, J.E. (1992) "*de novo* purine nucleotide biosynthesis", in: Progress in Nucleic Acid Research and Molecular Biology, vol. 42, Academic Press:, p. 259-287; and Michal, G. (1999) "Nucleotides and Nucleosides", Chapter 8 in: Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, Wiley: New York). Purine metabolism has been the subject of intensive research, and is essential to the normal functioning of the cell. Impaired purine metabolism in higher animals can cause severe disease, such as gout. Purine nucleotides are synthesized from

- 18 -

ribose-5-phosphate, in a series of steps through the intermediate compound inosine-5'-phosphate (IMP), resulting in the production of guanosine-5'-monophosphate (GMP) or adenosine-5'-monophosphate (AMP), from which the triphosphate forms utilized as nucleotides are readily formed. These compounds are also utilized as energy stores, so  
5 their degradation provides energy for many different biochemical processes in the cell. Pyrimidine biosynthesis proceeds by the formation of uridine-5'-monophosphate (UMP) from ribose-5-phosphate. UMP, in turn, is converted to cytidine-5'-triphosphate (CTP). The deoxy- forms of all of these nucleotides are produced in a one step reduction reaction from the diphosphate ribose form of the nucleotide to the diphosphate  
10 deoxyribose form of the nucleotide. Upon phosphorylation, these molecules are able to participate in DNA synthesis.

#### *D. Trehalose Metabolism and Uses*

Trehalose consists of two glucose molecules, bound in  $\alpha$ ,  $\alpha$ -1,1 linkage. It is  
15 commonly used in the food industry as a sweetener, an additive for dried or frozen foods, and in beverages. However, it also has applications in the pharmaceutical, cosmetics and biotechnology industries (see, for example, Nishimoto *et al.*, (1998) U.S. Patent No. 5,759,610; Singer, M.A. and Lindquist, S. (1998) *Trends Biotech.* 16: 460-467; Paiva, C.L.A. and Panek, A.D. (1996) *Biotech. Ann. Rev.* 2: 293-314; and  
20 Shiosaka, M. (1997) *J. Japan* 172: 97-102). Trehalose is produced by enzymes from many microorganisms and is naturally released into the surrounding medium, from which it can be collected using methods known in the art.

#### II. Maintenance of Homeostasis in *C. glutamicum* and Environmental Adaptation

25 The metabolic and other biochemical processes by which cells function are sensitive to environmental conditions such as temperature, pressure, solute concentration, and availability of oxygen. When one or more such environmental condition is perturbed or altered in a fashion that is incompatible with the normal functioning of these cellular processes, the cell must act to maintain an intracellular  
30 environment which will permit them to occur despite the hostile extracellular environment. Gram positive bacterial cells, such as *C. glutamicum* cells, have a number of mechanisms by which internal homeostasis may be maintained despite unfavorable

- 19 -

extracellular conditions. These include a cell wall, proteins which are able to degrade possibly toxic aromatic and aliphatic compounds, mechanisms of proteolysis whereby misfolded or misregulated proteins may be rapidly destroyed, and catalysts which permit intracellular reactions to occur which would not normally take place under the

5 conditions optimal for bacterial growth.

Aside from merely surviving in a hostile environment, bacterial cells (*e.g.* *C. glutamicum* cells) are also frequently able to adapt such that they are able to take advantage of such conditions. For example, cells in an environment lacking desired carbon sources may be able to adapt to growth on a less-suitable carbon source. Also,

10 cells may be able to utilize less desirable inorganic compounds when the commonly utilized ones are unavailable. *C. glutamicum* cells possess a number of genes which permit them to adapt to utilize inorganic and organic molecules which they would normally not encounter under optimal growth conditions as nutrients and precursors for metabolism. Aspects of cellular processes involved in homeostasis and adaptation are

15 further explicated below.

#### *A. Modification and Degradation of Aromatic and Aliphatic Compounds*

Bacterial cells are routinely exposed to a variety of aromatic and aliphatic compounds in nature. Aromatic compounds are organic molecules having a cyclic ring

20 structure, while aliphatic compounds are organic molecules having open chain structures rather than ring structures. Such compounds may arise as by-products of industrial processes (*e.g.*, benzene or toluene), but may also be produced by certain microorganisms (*e.g.*, alcohols). Many of these compounds are toxic to cells, particularly the aromatic compounds, which are highly reactive due to the high-energy ring structure. Thus, certain

25 bacteria have developed mechanisms by which they are able to modify or degrade these compounds such that they are no longer hazardous to the cell. Cells may possess enzymes that are able to, for example, hydroxylate, isomerize, or methylate aromatic or aliphatic compounds such that they are either rendered less toxic, or such that the modified form is able to be processed by standard cellular waste and degradation pathways. Also, cells may

30 possess enzymes which are able to specifically degrade one or more such potentially hazardous substance, thereby protecting the cell. Principles and examples of these types of modification and degradation processes in bacteria are described in several publications,



e.g., Sahm, H. (1999) "Procaryotes in Industrial Production" in Lengeler, J.W. *et al.*, eds. *Biology of the Procaryotes*, Thieme Verlag: Stuttgart; and Schlegel, H.G. (1992) *Allgemeine Mikrobiologie*, Thieme: Stuttgart).

Aside from simply inactivating hazardous aromatic or aliphatic compounds, many  
5 bacteria have evolved to be able to utilize these compounds as carbon sources for continued  
metabolism when the preferred carbon sources of the cell are not available. For example,  
Pseudomonas strains able to utilize toluene, benzene, and 1,10-dichlorodecane as carbon  
sources are known (Chang, B.V. *et al.* (1997) *Chemosphere* 35(12): 2807-2815; Wischnak,  
C. *et al.* (1998) *Appl. Environ. Microbiol.* 64(9): 3507-3511; Churchill, S.A. *et al.* (1999)  
10 *Appl. Environ. Microbiol.* 65(2): 549-552). There are similar examples from many other  
bacterial species which are known in the art.

The ability of certain bacteria to modify or degrade aromatic and aliphatic  
compounds has begun to be exploited. Petroleum is a complex mixture of chemicals which  
includes aliphatic molecules and aromatic compounds. By applying bacteria having the  
15 ability to degrade or modify these toxic compounds to an oil spill, for example, it is possible  
to eliminate much of the environmental damage with high efficiency and low cost (see, for  
example, Smith, M.R. (1990) "The biodegradation of aromatic hydrocarbons by bacteria"  
*Biodegradation* 1(2-3): 191-206; and Suyama, T. *et al.* (1998) "Bacterial isolates degrading  
aliphatic polycarbonates," *FEMS Microbiol. Lett.* 161(2): 255-261).

20

#### *B. Metabolism of Inorganic Compounds*

Cells (e.g., bacterial cells) contain large quantities of different molecules, such as  
water, inorganic ions, and organic substances (e.g., proteins, sugars, and other  
macromolecules). The bulk of the mass of a typical cell consists of only 4 types of atoms:  
25 carbon, oxygen, hydrogen, and nitrogen. Although they represent a smaller percentage of  
the content of a cell, inorganic substances are equally as important to the proper functioning  
of the cell. Such molecules include phosphorous, sulfur, calcium, magnesium, iron, zinc,  
manganese, copper, molybdenum, tungsten, and cobalt. Many of these compounds are  
critical for the construction of important molecules, such as nucleotides (phosphorous) and  
30 amino acids (nitrogen and sulfur). Others of these inorganic ions serve as cofactors for  
enzymic reactions or contribute to osmotic pressure. All such molecules must be taken up  
by the bacterium from the surrounding environment.

For each of these inorganic compounds it is desirable for the bacterium to take up the form which can be most readily used by the standard metabolic machinery of the cell. However, the bacterium may encounter environments in which these preferred forms are not readily available. In order to survive under these circumstances, it is important for bacteria  
5 to have additional biochemical mechanisms which are able to convert less metabolically active but readily available forms of these inorganic compounds to ones which may be used in cellular metabolism. Bacteria frequently possess a number of genes encoding enzymes for this purpose, which are not expressed unless the desired inorganic species are not available. Thus, these genes for the metabolism of various inorganic compounds serve as  
10 another tool which bacteria may use to adapt to suboptimal environmental conditions.

After carbon, the most important element in the cell is nitrogen. A typical bacterial cell contains between 12-15% nitrogen. It is a constituent of amino acids and nucleotides, as well as many other important molecules in the cell. Further, nitrogen may serve as a substitute for oxygen as a terminal electron acceptor in energy metabolism. Good sources  
15 of nitrogen include many organic and inorganic compounds, such ammonia gas or ammonia salts (*e.g.*,  $\text{NH}_4\text{Cl}$ ,  $(\text{NH}_4)_2\text{SO}_4$ , or  $\text{NH}_4\text{OH}$ ), nitrates, urea, amino acids, or complex nitrogen sources like corn steep liquor, soy bean flour, soy bean protein, yeast extract, meat extract, etc. Ammonia nitrogen is fixed by the action of particular enzymes: glutamate dehydrogenase, glutamine synthase, and glutamine-2-oxoglutarate aminotransferase. The  
20 transfer of amino-nitrogen from one organic molecule to another is accomplished by the aminotransferases, a class of enzymes which transfer one amino group from an alpha-amino acid to an alpha-keto acid. Nitrate may be reduced via nitrate reductase, nitrite reductase, and further redox enzymes until it is converted to molecular nitrogen or ammonia, which may be readily utilized by the cell in standard metabolic pathways.

25 Phosphorous is typically found intracellularly in both organic and inorganic forms, and may be taken up by the cell in either of these forms as well, though most microorganisms preferentially take up inorganic phosphate. The conversion of organic phosphate to a form which the cell can utilize requires the action of phosphatases (*e.g.*, phytases, which hydrolyze phytate-yielding phosphate and inositol derivatives). Phosphate  
30 is a key element in the synthesis of nucleic acids, and also has a significant role in cellular energy metabolism (*e.g.*, in the synthesis of ATP, ADP, and AMP).

- 22 -

Sulfur is a requirement for the synthesis of amino acids (*e.g.*, methionine and cysteine), vitamins (*e.g.*, thiamine, biotin, and lipoic acid) and iron sulfur proteins. Bacteria obtain sulfur primarily from inorganic sulfate, though thiosulfate, sulfite, and sulfide are also commonly utilized. Under conditions where these compounds may not be readily available, many bacteria express genes which enable them to utilize sulfonate compounds such as 2-aminosulfonate (taurine) (Kertesz, M.A. (1993) "Proteins induced by sulfate limitation in *Escherichia coli*, *Pseudomonas putida*, or *Staphylococcus aureus*." *J. Bacteriol.* 175: 1187-1190).

Other inorganic atoms, *e.g.*, metal or calcium ions, are also critical for the viability of cells. Iron, for example, plays a key role in redox reactions and is a cofactor of iron-sulfur proteins, heme proteins, and cytochromes. The uptake of iron into bacterial cells may be accomplished by the action of siderophores, chelating agents which bind extracellular iron ions and translocate them to the interior of the cell. For reference on the metabolism of iron and other inorganic compounds, see: Lengeler *et al.* (1999) *Biology of Prokaryotes*, Thieme Verlag: Stuttgart; Neidhardt, F.C. *et al.*, eds. *Escherichia coli* and *Salmonella*. ASM Press: Washington, D.C.; Sonenshein, A.L. *et al.*, eds. (1997) *Bacillus subtilis* and Other Gram-Positive Bacteria, ASM Press: Washington, D.C.; Voet, D. and Voet, J.G. (1992) *Biochemie*, VCH: Weinheim; Brock, T.D. and Madigan, M.T. (1991) *Biology of Microorganisms*, 6<sup>th</sup> ed. Prentice Hall: Englewood Cliffs, p. 267-269; Rhodes, P.M. and Stanbury, P.F. *Applied Microbial Physiology – A Practical Approach*, Oxford Univ. Press: Oxford.

### C. Enzymes and Proteolysis

The intracellular conditions for which bacteria such as *C. glutamicum* are optimized are frequently not conditions under which many biochemical reactions would normally take place. In order to make such reactions proceed under physiological conditions, cells utilize enzymes. Enzymes are proteinaceous biological catalysts, spatially orienting reacting molecules or providing a specialized environment such that the energy barrier to a biochemical reaction is lowered. Different enzymes catalyze different reactions, and each enzyme may be the subject of transcriptional, translational, or posttranslational regulation such that the reaction will only take place under appropriate conditions and at specified times. Enzymes may contribute to the

- 23 -

degradation (*e.g.*, the proteases), synthesis (*e.g.*, the synthases), or modification (*e.g.*, transferases or isomerases) of compounds, all of which enable the production of necessary compounds within the cell. This, in turn, contributes to the maintenance of cellular homeostasis.

5           However, the fact that enzymes are optimized for activity under the physiological conditions at which the bacterium is most viable means that when environmental conditions are perturbed, there is a significant possibility that enzyme activity will also be perturbed. For example, changes in temperature may result in aberrantly folded proteins, and the same is true for changes of pH – protein folding is  
10 largely dependent on electrostatic and hydrophobic interactions of amino acids within the polypeptide chain, so any alteration to the charges on individual amino acids (as might be brought about by a change in cellular pH) may have a profound effect on the ability of the protein to correctly fold. Changes in temperature effectively change the amount of kinetic energy that the polypeptide molecule possesses, which affects the  
15 ability of the polypeptide to settle into a correctly folded, energetically stable configuration. Misfolded proteins may be harmful to the cell for two reasons. First, the aberrantly folded protein may have a similarly aberrant activity, or no activity whatsoever. Second, misfolded proteins may lack the conformational regions necessary for proper regulation by other cellular systems and thus may continue to be active but in  
20 an uncontrolled fashion.

          The cell has a mechanism by which misfolded enzymes and regulatory proteins may be rapidly destroyed before any damage occurs to the cell: proteolysis. Proteins such as those of the *la/lon* family and those of the *Clp* family specifically recognize and degrade misfolded proteins (see, *e.g.*, Sherman, M.Y., Goldberg, A.L. (1999) *EXS* 77:  
25 57-78 and references therein and Porankiewicz J. (1999) *Molec. Microbiol.* 32(3): 449-58, and references therein; Neidhardt, F.C., *et al.* (1996) *E. coli* and *Salmonella*, ASM Press: Washington, D.C. and references therein; and Pritchard, G.G., and Coolbear, T. (1993) *FEMS Microbiol. Rev.* 12(1-3): 179-206 and references therein). These enzymes bind to misfolded or unfolded proteins and degrade them in an ATP-dependent manner.  
30 Proteolysis thus serves as an important mechanism employed by the cell to prevent damage to normal cellular functions upon environmental changes, and it further permits

- 24 -

cells to survive under conditions and in environments which would otherwise be toxic due to misregulated and/or aberrant enzyme or regulatory activity.

Proteolysis also has important functions in the cell under optimal environmental conditions. Within normal metabolic processes, proteases aid in the hydrolysis of peptide bonds, in the catabolism of complex molecules to provide necessary degradation products, and in protein modification. Secreted proteases play an important role in the catabolism of external nutrients even prior to the entry of these compounds into the cell. Further, proteolytic activity itself may serve regulatory functions; sporulation in *B. subtilis* and cell cycle progression in *Caulobacter* spp. are known to be regulated by key proteolytic events in each of these species (Gottesman, S. (1999) *Curr. Opin. Microbiol.* 2(2): 142-147). Thus, proteolytic processes are key for cellular survival under both suboptimal and optimal environmental conditions, and contribute to the overall maintenance of homeostasis in cells.

#### 15 D. Cell Wall Production and Rearrangements

While the biochemical machinery of the cell may be able to readily adapt to different and possibly unfavorable environments, cells still require a general mechanism by which they may be protected from the environment. For many bacteria, the cell wall affords such protection, and also plays roles in adhesion, cell growth and division, and transport of desired solutes and waste materials.

In order to function, cells require intracellular concentrations of metabolites and other molecules that are substantially higher than those of the surrounding media. Since these metabolites are largely prevented from leaving the cell due to the presence of the hydrophobic membrane, the tendency of the system is for water molecules to enter the cell from the external medium such that the interior concentrations of solutes match the exterior concentrations. Water molecules are readily able to cross the cellular membrane, and this membrane is not able to withstand the resulting swelling and pressure, which may lead to osmotic lysis of the cell. The rigidity of the cell wall greatly improves the ability of the cell to tolerate these pressures, and offers a further barrier to the unwanted diffusion of these metabolites and desired solutes from the cell. Similarly, the cell wall also serves to prevent unwanted material from entering the cell.

The cell wall also participates in a number of other cellular processes, such as adhesion and cell growth and division. Due to the fact that the cell wall completely surrounds the cell, any interaction of the cell with its surroundings must be mediated by the cell wall. Thus, the cell wall must participate in any adherence of the cell to other cells and to desired surfaces. Further, the cell cannot grow or divide without concomitant changes in the cell wall. Since the protection that the wall affords requires its presence during growth, morphogenesis and multiplication, one of the key steps in cell division is cell wall synthesis within the cell such that a new cell divides from the old. Thus, frequently cell wall biosynthesis is regulated in tandem with cell growth and cell division (see, *e.g.*, Sonenshein, A.L. et al, eds. (1993) *Bacillus subtilis* and Other Gram-Positive Bacteria, ASM: Washington, D.C.).

The structure of the cell wall varies between gram-positive and gram-negative bacteria. However, in both types, the fundamental structural unit of the wall remains similar: an overlapping lattice of two polysaccharides, N-acetyl glucosamine (NAG) and N-acetyl muramic acid (NAM) which are cross-linked by amino acids (most commonly L-alanine, D-glutamate, diaminopimelic acid, and D-alanine), termed 'peptidoglycan'. The processes involved in the synthesis of the cell wall are known (see, *e.g.*, Michal, G., ed. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, Wiley: New York).

In gram-negative bacteria, the inner cellular membrane is coated by a single-layered peptidoglycan (approximately 10 nm thick), termed the murein-sacculus. This peptidoglycan structure is very rigid, and its structure determines the shape of the organism. The outer surface of the murein-sacculus is covered with an outer membrane, containing porins and other membrane proteins, phospholipids, and lipopolysaccharides. To maintain a tight association with the outer membrane, the gram-negative cell wall also has interspersed lipid molecules which serve to anchor it to the surrounding membrane.

In gram-positive bacteria, such as *Corynebacterium glutamicum*, the cytoplasmic membrane is covered by a multi-layered peptidoglycan, which ranges from 20-80 nm in thickness (see, *e.g.*, Lengeler *et al.* (1999) *Biology of Prokaryotes* Thieme Verlag: Stuttgart, p. 913-918, p. 875-899, and p. 88-109 and references therein). The gram-positive cell wall also contains teichoic acid, a polymer of glycerol or ribitol linked through phosphate groups. Teichoic acid is also able to associate with amino acids, and forms covalent bonds with

muramic acid. Also present in the cell wall may be lipoteichoic acids and teichuronic acids. If present, cellular surface structures such as flagella or capsules will be anchored in this layer as well.

5 III. Elements and Methods of the Invention

The present invention is based, at least in part, on the discovery of novel molecules, referred to herein as HA nucleic acid and protein molecules, which participate in the maintenance of homeostasis in *C. glutamicum*, or which perform a function involved in the adaptation of this microorganism to different environmental  
10 conditions. In one embodiment, the HA molecules participate in *C. glutamicum* cell wall biosynthesis or rearrangements, in the metabolism of inorganic compounds, in the modification or degradation of aromatic or aliphatic compounds, or have an enzymatic or proteolytic activity. In a preferred embodiment, the activity of the HA molecules of the present invention with regard to *C. glutamicum* cell wall biosynthesis or  
15 rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or enzymatic or proteolytic activity has an impact on the production of a desired fine chemical by this organism. In a particularly preferred embodiment, the HA molecules of the invention are modulated in activity, such that the *C. glutamicum* cellular processes in which the HA molecules participate (*e.g.*, *C.*  
20 *glutamicum* cell wall biosynthesis or rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or enzymatic or proteolytic activity) are also altered in activity, resulting either directly or indirectly in a modulation of the yield, production, and/or efficiency of production of a desired fine chemical by *C. glutamicum*.

25 The language, "HA protein" or "HA polypeptide" includes proteins which participate in a number of cellular processes related to *C. glutamicum* homeostasis or the ability of *C. glutamicum* cells to adapt to unfavorable environmental conditions. For example, an HA protein may be involved in *C. glutamicum* cell wall biosynthesis or rearrangements, in the metabolism of inorganic compounds in *C. glutamicum*, in the  
30 modification or degradation of aromatic or aliphatic compounds in *C. glutamicum*, or have a *C. glutamicum* enzymatic or proteolytic activity. Examples of HA proteins include those encoded by the HA genes set forth in Table 1 and by the odd-numbered

SEQ ID NOs. The terms "HA gene" or "HA nucleic acid sequence" include nucleic acid sequences encoding an HA protein, which consist of a coding region and also corresponding untranslated 5' and 3' sequence regions. Examples of HA genes include those set forth in Table 1. The terms "production" or "productivity" are art-recognized and include the concentration of the fermentation product (for example, the desired fine chemical) formed within a given time and a given fermentation volume (*e.g.*, kg product per hour per liter). The term "efficiency of production" includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular rate of output of a fine chemical). The term "yield" or "product/carbon yield" is art-recognized and includes the efficiency of the conversion of the carbon source into the product (*i.e.*, fine chemical). This is generally written as, for example, kg product per kg carbon source. By increasing the yield or production of the compound, the quantity of recovered molecules, or of useful recovered molecules of that compound in a given amount of culture over a given amount of time is increased. The terms "biosynthesis" or a "biosynthetic pathway" are art-recognized and include the synthesis of a compound, preferably an organic compound, by a cell from intermediate compounds in what may be a multistep and highly regulated process. The terms "degradation" or a "degradation pathway" are art-recognized and include the breakdown of a compound, preferably an organic compound, by a cell to degradation products (generally speaking, smaller or less complex molecules) in what may be a multistep and highly regulated process. The language "metabolism" is art-recognized and includes the totality of the biochemical reactions that take place in an organism. The metabolism of a particular compound, then, (*e.g.*, the metabolism of an amino acid such as glycine) comprises the overall biosynthetic, modification, and degradation pathways in the cell related to this compound. The term "homeostasis" is art-recognized and includes all of the mechanisms utilized by a cell to maintain a constant intracellular environment despite the prevailing extracellular environmental conditions. A non-limiting example of such processes is the utilization of a cell wall to prevent osmotic lysis due to high intracellular solute concentrations. The term "adaptation" or "adaptation to an environmental condition" is art-recognized and includes mechanisms utilized by the cell to render the cell able to survive under nonpreferred environmental conditions (generally speaking, those environmental conditions in which one or more



5 favored nutrients are absent, or in which an environmental condition such as temperature, pH, osmolarity, oxygen percentage and the like fall outside of the optimal survival range of the cell). Many cells, including *C. glutamicum* cells, possess genes encoding proteins which are expressed under such environmental conditions and which permit continued growth in such suboptimal conditions.

10 In another embodiment, the HA molecules of the invention are capable of modulating the production of a desired molecule, such as a fine chemical, in a microorganism such as *C. glutamicum*. There are a number of mechanisms by which the alteration of an HA protein of the invention may directly affect the yield, production, and/or efficiency of production of a fine chemical from a *C. glutamicum* strain incorporating such an altered protein. For example, by engineering enzymes which modify or degrade aromatic or aliphatic compounds such that these enzymes are increased or decreased in activity or number, it may be possible to modulate the production of one or more fine chemicals which are the modification or degradation products of these compounds. Similarly, enzymes involved in the metabolism of inorganic compounds provide key molecules (e.g. phosphorous, sulfur, and nitrogen molecules) for the biosynthesis of such fine chemicals as amino acids, vitamins, and nucleic acids. By altering the activity or number of these enzymes in *C. glutamicum*, it may be possible to increase the conversion of these inorganic compounds (or to use alternate inorganic compounds) to thus permit improved rates of incorporation of inorganic atoms into these fine chemicals. Genetic engineering of *C. glutamicum* enzymes involved in general cellular processes may also directly improve fine chemical production, since many of these enzymes directly modify fine chemicals (e.g., amino acids) or the enzymes which are involved in fine chemical synthesis or secretion.

25 Modulation of the activity or number of cellular proteases may also have a direct effect on fine chemical production, since many proteases may degrade fine chemicals or enzymes involved in fine chemical production or breakdown.

30 Further, the aforementioned enzymes which participate in aromatic/aliphatic compound modification or degradation, general biocatalysis, inorganic compound metabolism or proteolysis are each themselves fine chemicals, desirable for their activity in various *in vitro* industrial applications. By altering the number of copies of the gene for one or more of these enzymes in *C. glutamicum* it may be possible to increase the

number of these proteins produced by the cell, thereby increasing the potential yield or efficiency of production of these proteins from large-scale *C. glutamicum* or related bacterial cultures.

The alteration of an HA protein of the invention may also indirectly affect the yield, production, and/or efficiency of production of a fine chemical from a *C. glutamicum* strain incorporating such an altered protein. For example, by modulating the activity and/or number of those proteins involved in the construction or rearrangement of the cell wall, it may be possible to modify the structure of the cell wall itself such that the cell is able to better withstand the mechanical and other stresses present during large-scale fermentative culture. Also, large-scale growth of *C. glutamicum* requires significant cell wall production. Modulation of the activity or number of cell wall biosynthetic or degradative enzymes may allow more rapid rates of cell wall biosynthesis, which in turn may permit increased growth rates of this microorganism in culture and thereby increase the number of cells producing the desired fine chemical.

By modifying the HA enzymes of the invention, one may also indirectly impact the yield, production, or efficiency of production of one or more fine chemicals from *C. glutamicum*. For example, many of the general enzymes in *C. glutamicum* may have a significant impact on global cellular processes (e.g., regulatory processes) which in turn have a significant effect on fine chemical metabolism. Similarly, proteases, enzymes which modify or degrade possibly toxic aromatic or aliphatic compounds, and enzymes which promote the metabolism of inorganic compounds all serve to increase the viability of *C. glutamicum*. The proteases aid in the selective removal of misfolded or misregulated proteins, such as those that might occur under the relatively stressful environmental conditions encountered during large-scale fermentor culture. By altering these proteins, it may be possible to further enhance this activity and to improve the viability of *C. glutamicum* in culture. The aromatic/aliphatic modification or degradation proteins not only serve to detoxify these waste compounds (which may be encountered as impurities in culture medium or as waste products from cells themselves), but also to permit the cells to utilize alternate carbon sources if the optimal carbon source is limiting in the culture. By increasing their number and/or activity, the survival of *C. glutamicum* cells in culture may be enhanced. The inorganic metabolism

- 30 -

proteins of the invention supply the cell with inorganic molecules required for all protein and nucleotide (among others) synthesis, and thus are critical for the overall viability of the cell. An increase in the number of viable cells producing one or more desired fine chemicals in large-scale culture should result in a concomitant increase in the yield, production, and/or efficiency of production of the fine chemical in the culture.

The isolated nucleic acid sequences of the invention are contained within the genome of a *Corynebacterium glutamicum* strain available through the American Type Culture Collection, given designation ATCC 13032. The nucleotide sequence of the isolated *C. glutamicum* HA DNAs and the predicted amino acid sequences of the *C. glutamicum* HA proteins are shown in the Sequence Listing as odd-numbered SEQ ID NOs and even-numbered SEQ ID NOs, respectively, respectively. Computational analyses were performed which classified and/or identified these nucleotide sequences as sequences which encode proteins that participate in *C. glutamicum* cell wall biosynthesis or rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or that have a *C. glutamicum* enzymatic or proteolytic activity.

The present invention also pertains to proteins which have an amino acid sequence which is substantially homologous to an amino acid sequence of the invention (*e.g.*, the sequence of an even-numbered SEQ ID NO of the Sequence Listing). As used herein, a protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence is least about 50% homologous to the selected amino acid sequence, *e.g.*, the entire selected amino acid sequence. A protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence can also be least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, or 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to the selected amino acid sequence.

The HA protein or a biologically active portion or fragment thereof of the invention can participate in the maintenance of homeostasis in *C. glutamicum*, or can perform a function involved in the adaptation of this microorganism to different environmental conditions, or have one or more of the activities set forth in Table 1.

Various aspects of the invention are described in further detail in the following subsections.

### A. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode HA polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of HA-encoding nucleic acid (e.g., HA DNA). As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses untranslated sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 100 nucleotides of sequence upstream from the 5' end of the coding region and at least about 20 nucleotides of sequence downstream from the 3' end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated HA nucleic acid molecule can contain less than about 5 kb, 4kb, 3kb, 2kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g., a *C. glutamicum* cell). Moreover, an "isolated" nucleic acid molecule, such as a DNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having a nucleotide sequence of an odd-numbered SEQ ID NO of the Sequence Listing, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, a *C. glutamicum* HA DNA can be isolated from a *C. glutamicum* library using all or portion of one of the odd-numbered SEQ ID NO sequences of the Sequence Listing as a hybridization probe and standard hybridization techniques (e.g., as described in Sambrook, J., Fritsch, E. F., and Maniatis,

- 32 -

T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Moreover, a nucleic acid molecule encompassing all or a portion of one of the nucleic acid sequences of the invention (e.g., an odd-numbered SEQ ID NO:) can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this sequence (e.g., a nucleic acid molecule encompassing all or a portion of one of the nucleic acid sequences of the invention (e.g., an odd-numbered SEQ ID NO of the Sequence Listing) can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this same sequence). For example, mRNA can be isolated from normal endothelial cells (e.g., by the guanidinium-thiocyanate extraction procedure of Chirgwin *et al.* (1979) *Biochemistry* 18: 5294-5299) and DNA can be prepared using reverse transcriptase (e.g., Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL). Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed based upon one of the nucleotide sequences shown in the Sequence Listing. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an HA nucleotide sequence can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises one of the nucleotide sequences shown in the Sequence Listing. The nucleic acid sequences of the invention, as set forth in the Sequence Listing, correspond to the *Corynebacterium glutamicum* HA DNAs of the invention. This DNA comprises sequences encoding HA proteins (i.e., the "coding region", indicated in each odd-numbered SEQ ID NO: sequence in the Sequence Listing), as well as 5' untranslated sequences and 3' untranslated sequences, also indicated in each odd-numbered SEQ ID NO: in the Sequence Listing. Alternatively, the nucleic acid molecule can comprise only the coding region of any of the sequences in nucleic acid sequences of the Sequence Listing

For the purposes of this application, it will be understood that each of the nucleic acid and amino acid sequences set forth in the Sequence Listing has an identifying RXA, RXN, RXS, or RXC number having the designation "RXA," "RXN," "RXS, or "RXC" followed by 5 digits (*i.e.*, RXA02458, RXN00249, RXS00153, or RXC00963).

5 Each of the nucleic acid sequences comprises up to three parts: a 5' upstream region, a coding region, and a downstream region. Each of these three regions is identified by the same RXA, RXN, RXS, or RXC designation to eliminate confusion. The recitation "one of the odd-numbered sequences in of the Sequence Listing", then, refers to any of the nucleic acid sequences in the Sequence Listing, which may also be distinguished by

10 their differing RXA, RXN, RXS, or RXC designations. The coding region of each of these sequences is translated into a corresponding amino acid sequence, which is also set forth in the Sequence Listing, as an even-numbered SEQ ID NO: immediately following the corresponding nucleic acid sequence . For example, the coding region for RXA02548 is set forth in SEQ ID NO:1, while the amino acid sequence which it

15 encodes is set forth as SEQ ID NO:2. The sequences of the nucleic acid molecules of the invention are identified by the same RXA, RXN, RXS, or RXC designations as the amino acid molecules which they encode, such that they can be readily correlated. For example, the amino acid sequences designated RXA02458, RXN00249, RXS00153, and RXC00963 are translations of the coding regions of the nucleotide sequences of nucleic

20 acid molecules RXA02458, RXN00249, RXS00153, and RXC00963, respectively. of the correspondence between the RXA, RXN, RXS, and RXC nucleotide and amino acid sequences of the invention and their assigned SEQ ID NOs is set forth in Table 1.

Several of the genes of the invention are "F-designated genes". An F-designated gene includes those genes set forth in Table 1 which have an 'F' in front of the RXA, RXN, RXS, or RXC designation. For example, SEQ ID NO:5, designated, as indicated

25 on Table 1, as "F RXA00249", is an F-designated gene, as are SEQ ID NOs: 11, 15, and 33 (designated on Table 1 as "F RXA02264", "F RXA02274", and "F RXA00675", respectively).

In one embodiment, the nucleic acid molecules of the present invention are not

30 intended to include those compiled in Table 2. In the case of the *dapD* gene, a sequence for this gene was published in Wehrmann, A., *et al.* (1998) *J. Bacteriol.* 180(12): 3159-3165. However, the sequence obtained by the inventors of the present application is

significantly longer than the published version. It is believed that the published version relied on an incorrect start codon, and thus represents only a fragment of the actual coding region.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of one of the nucleotide sequences of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the Sequence Listing), or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences of the invention is one which is sufficiently complementary to one of the nucleotide sequences the Sequence Listing (*e.g.*, the sequence of an odd-numbered SEQ ID NO:) such that it can hybridize to one of the nucleotide sequences of the invention, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence of the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the Sequence Listing), or a portion thereof. Ranges and identity values intermediate to the above-recited ranges, (*e.g.*, 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. In an additional preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to one of the nucleotide sequences of the invention, or a portion thereof.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of the sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of an HA protein. The nucleotide sequences determined from the cloning of the HA genes from *C. glutamicum* allows for the generation of probes and primers designed for use in identifying and/or cloning HA

- 35 -

homologues in other cell types and organisms, as well as HA homologues from other *Corynebacteria* or related species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50 or 75 consecutive nucleotides of a sense strand of one of the nucleotide sequences of the invention (*e.g.*, a sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing), an anti-sense sequence of one of these sequences, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of the invention can be used in PCR reactions to clone HA homologues.

Probes based on the HA nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, *e.g.* the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells which misexpress an HA protein, such as by measuring a level of an HA-encoding nucleic acid in a sample of cells, *e.g.*, detecting HA mRNA levels or determining whether a genomic HA gene has been mutated or deleted.

In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO of the Sequence Listing) such that the protein or portion thereof maintains the ability to participate in the maintenance of homeostasis in *C. glutamicum*, or to perform a function involved in the adaptation of this microorganism to different environmental conditions. As used herein, the language "sufficiently homologous" refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (*e.g.*, an amino acid residue which has a similar side chain as an amino acid residue in a sequence of one of the even-numbered SEQ ID NOs of the Sequence Listing) amino acid residues to an amino acid sequence of the invention such that the protein or portion thereof is able to participate in the maintenance of homeostasis in *C. glutamicum*, or to perform a function involved in the adaptation of this microorganism to different environmental conditions. Proteins involved in *C. glutamicum* cell wall biosynthesis or rearrangements, metabolism of



- 36 -

inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or that have a *C. glutamicum* enzymatic or proteolytic activity, as described herein, may play a role in the production and secretion of one or more fine chemicals. Examples of such activities are also described herein. Thus, "the function of an HA protein"

5 contributes either directly or indirectly to the yield, production, and/or efficiency of production of one or more fine chemicals. Examples of HA protein activities are set forth in Table 1.

In another embodiment, the protein is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, 90-95%, and most  
10 preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of the invention (e.g., a sequence of an even-numbered SEQ ID NO: of the Sequence Listing).

Portions of proteins encoded by the HA nucleic acid molecules of the invention are preferably biologically active portions of one of the HA proteins. As used herein,  
15 the term "biologically active portion of an HA protein" is intended to include a portion, e.g., a domain/motif, of an HA protein that can participate in the maintenance of homeostasis in *C. glutamicum*, or that can perform a function involved in the adaptation of this microorganism to different environmental conditions, or has an activity as set forth in Table 1. To determine whether an HA protein or a biologically active portion  
20 thereof can participate in *C. glutamicum* cell wall biosynthesis or rearrangements, metabolism of inorganic compounds, modification or degradation of aromatic or aliphatic compounds, or has a *C. glutamicum* enzymatic or proteolytic activity, an assay of enzymatic activity may be performed. Such assay methods are well known to those of ordinary skill in the art, as detailed in Example 8 of the Exemplification.

25 Additional nucleic acid fragments encoding biologically active portions of an HA protein can be prepared by isolating a portion of one of the amino acid sequences in of the invention (e.g., a sequence of an even-numbered SEQ ID NO: of the Sequence Listing), expressing the encoded portion of the HA protein or peptide (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the  
30 HA protein or peptide.

The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences of the invention (e.g., a sequence of an odd-numbered SEQ ID

- 37 -

NO: of the Sequence Listing) (and portions thereof) due to degeneracy of the genetic code and thus encode the same HA protein as that encoded by the nucleotide sequences shown in of the invention. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid  
5 sequence shown in the Sequence Listing (e.g., an even-numbered SEQ ID NO:). In a still further embodiment, the nucleic acid molecule of the invention encodes a full length *C. glutamicum* protein which is substantially homologous to an amino acid sequence of the invention (encoded by an open reading frame shown in an odd-numbered SEQ ID NO: of the Sequence Listing).

10 It will be understood by one of ordinary skill in the art that in one embodiment the sequences of the invention are not meant to include the sequences of the prior art, such as those Genbank sequences set forth in Tables 2 or 4 which were available prior to the present invention. In one embodiment, the invention includes nucleotide and amino acid sequences having a percent identity to a nucleotide or amino acid sequence of the  
15 invention which is greater than that of a sequence of the prior art (e.g., a Genbank sequence (or the protein encoded by such a sequence) set forth in Tables 2 or 4). For example, the invention includes a nucleotide sequence which is greater than and/or at least 39% identical to the nucleotide sequence designated RXA00471 (SEQ ID NO:293), a nucleotide sequence which is greater than and/or at least 41% identical to the  
20 nucleotide sequence designated RXA00500 (SEQ ID NO:143), and a nucleotide sequence which is greater than and/or at least 35% identical to the nucleotide sequence designated RXA00502 (SEQ ID NO:147). One of ordinary skill in the art would be able to calculate the lower threshold of percent identity for any given sequence of the invention by examining the GAP-calculated percent identity scores set forth in Table 4  
25 for each of the three top hits for the given sequence, and by subtracting the highest GAP-calculated percent identity from 100 percent. One of ordinary skill in the art will also appreciate that nucleic acid and amino acid sequences having percent identities greater than the lower threshold so calculated (e.g., at least 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%,  
30 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%,

88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more identical) are also encompassed by the invention.

In addition to the *C. glutamicum* HA nucleotide set forth in the Sequence Listing as odd-numbered SEQ ID NOs, it will be appreciated by those of ordinary skill in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of HA proteins may exist within a population (e.g., the *C. glutamicum* population). Such genetic polymorphism in the HA gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding an HA protein, preferably a *C. glutamicum* HA protein. Such natural variations can typically result in 1-5% variance in the nucleotide sequence of the HA gene. Any and all such nucleotide variations and resulting amino acid polymorphisms in HA that are the result of natural variation and that do not alter the functional activity of HA proteins are intended to be within the scope of the invention.

Nucleic acid molecules corresponding to natural variants and non-*C. glutamicum* homologues of the *C. glutamicum* HA DNA of the invention can be isolated based on their homology to the *C. glutamicum* HA nucleic acid disclosed herein using the *C. glutamicum* DNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising a nucleotide sequence of an odd-numbered SEQ ID NO: of the Sequence Listing. In other embodiments, the nucleic acid is at least 30, 50, 100, 250 or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those of ordinary skill in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are

- 39 -

hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a nucleotide sequence of the invention corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (*e.g.*, encodes a natural protein). In one embodiment, the nucleic acid encodes a natural *C. glutamicum* HA protein.

In addition to naturally-occurring variants of the HA sequence that may exist in the population, one of ordinary skill in the art will further appreciate that changes can be introduced by mutation into a nucleotide sequence of the invention, thereby leading to changes in the amino acid sequence of the encoded HA protein, without altering the functional ability of the HA protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in a nucleotide sequence of the invention. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of one of the HA proteins (*e.g.*, an even-numbered SEQ ID NO: of the Sequence Listing) without altering the activity of said HA protein, whereas an "essential" amino acid residue is required for HA protein activity. Other amino acid residues, however, (*e.g.*, those that are not conserved or only semi-conserved in the domain having HA activity) may not be essential for activity and thus are likely to be amenable to alteration without altering HA activity.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding HA proteins that contain changes in amino acid residues that are not essential for HA activity. Such HA proteins differ in amino acid sequence from a sequence of an even-numbered SEQ ID NO: of the Sequence Listing yet retain at least one of the HA activities described herein. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 50% homologous to an amino acid sequence of the invention and is capable of participating in the maintenance of homeostasis in *C. glutamicum*, or of performing a function involved in the adaptation of this microorganism to different environmental conditions, or has one or more of the activities set forth in Table 1. Preferably, the protein encoded by the nucleic acid

- 40 -

molecule is at least about 50-60% homologous to the amino acid sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing, more preferably at least about 60-70% homologous to one of these sequences in, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to one of these sequences, and most preferably at least about 96%, 97%, 98%, or 99% homologous to one of the amino acid sequences of the invention

To determine the percent homology of two amino acid sequences (*e.g.*, one of the amino acid sequences of the invention and a mutant form thereof) or of two nucleic acids, the sequences are aligned for optimal comparison purposes (*e.g.*, gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in one sequence (*e.g.*, one of the amino acid sequences of the invention) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (*e.g.*, a mutant form of the amino acid sequence), then the molecules are homologous at that position (*i.e.*, as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity"). The percent homology between the two sequences is a function of the number of identical positions shared by the sequences (*i.e.*, % homology = # of identical positions/total # of positions x 100).

An isolated nucleic acid molecule encoding an HA protein homologous to a protein sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) can be created by introducing one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of the invention such that one or more amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced into one of the nucleotide sequences of the invention by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (*e.g.*, lysine, arginine, histidine), acidic side chains (*e.g.*, aspartic acid, glutamic

- 41 -

acid), uncharged polar side chains (*e.g.*, glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (*e.g.*, alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (*e.g.*, threonine, valine, isoleucine) and aromatic side chains (*e.g.*, tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an HA protein is preferably replaced with another amino acid residue from the same side chain family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an HA coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an HA activity described herein to identify mutants that retain HA activity. Following mutagenesis of the nucleotide sequence of one of the odd-numbered SEQ ID NOs of the Sequence Listing, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using, for example, assays described herein (see Example 8 of the Exemplification).

In addition to the nucleic acid molecules encoding HA proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, *e.g.*, complementary to the coding strand of a double-stranded DNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire HA coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an HA protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues (*e.g.*, the entire coding region of SEQ ID NO. 3 (RXN00249) comprises nucleotides 1 to 957). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding HA. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (*i.e.*, also referred to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding HA disclosed herein (*e.g.*, the sequences set forth as odd-numbered SEQ ID NOs in the Sequence Listing), antisense

nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of HA mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of HA mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of HA mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (*e.g.*, an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, *e.g.*, phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (*i.e.*, RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a cell or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an HA protein to thereby inhibit expression of the protein, *e.g.*, by inhibiting transcription and/or translation. The hybridization can be by  
5 conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, *e.g.*, by linking the antisense nucleic acid molecule to a peptide or  
10 an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong prokaryotic, viral, or eukaryotic promoter are preferred.

15 In yet another embodiment, the antisense nucleic acid molecule of the invention is an  $\alpha$ -anomeric nucleic acid molecule. An  $\alpha$ -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual  $\beta$ -units, the strands run parallel to each other (Gaultier *et al.* (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-*o*-  
20 methylribonucleotide (Inoue *et al.* (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue *et al.* (1987) *FEBS Lett.* 215:327-330).

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they  
25 have a complementary region. Thus, ribozymes (*e.g.*, hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave HA mRNA transcripts to thereby inhibit translation of HA mRNA. A ribozyme having specificity for an HA-encoding nucleic acid can be designed based upon the nucleotide sequence of an HA DNA molecule disclosed herein (*i.e.*, SEQ ID  
30 NO. 3 (RXN00249)). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an HA-encoding mRNA. See, *e.g.*, Cech *et al.*



U.S. Patent No. 4,987,071 and Cech *et al.* U.S. Patent No. 5,116,742. Alternatively, HA mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, *e.g.*, Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

- 5           Alternatively, HA gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of an HA nucleotide sequence (*e.g.*, an HA promoter and/or enhancers) to form triple helical structures that prevent transcription of an HA gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. *et al.* (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and
- 10   Maher, L.J. (1992) *Bioassays* 14(12):807-15.

#### *B. Recombinant Expression Vectors and Host Cells*

- Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an HA protein (or a portion thereof). As
- 15   used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of
- 20   autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to
- 25   which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector.
- 30   However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which are

5 operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (*e.g.*, in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory

10 sequence" is intended to include promoters, enhancers and other expression control elements (*e.g.*, polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and

15 those which direct expression of the nucleotide sequence only in certain host cells. Preferred regulatory sequences are, for example, promoters such as *cos*-, *tac*-, *trp*-, *tet*-, *trp-tet*-, *lpp*-, *lac*-, *lpp-lac*-, *lacI<sup>q</sup>*-, *T7*-, *T5*-, *T3*-, *gal*-, *trc*-, *ara*-, *SP6*-, *arny*, *SPO2*,  $\lambda$ -*P<sub>R</sub>*- or  $\lambda$  *P<sub>L</sub>*, which are used preferably in bacteria. Additional regulatory sequences are, for example, promoters from yeasts and fungi, such as *ADC1*, *MF $\alpha$* , *AC*, *P-60*, *CYC1*,

20 *GAPDH*, *TEF*, *rp28*, *ADH*, promoters from plants such as *CaMV/35S*, *SSU*, *OCS*, *lib4*, *usp*, *STLS1*, *B33*, *nos* or *ubiquitin*- or *phaseolin*-promoters. It is also possible to use artificial promoters. It will be appreciated by those of ordinary skill in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors

25 of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein (*e.g.*, HA proteins, mutant forms of HA proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of HA proteins in prokaryotic or eukaryotic cells. For example, HA genes

30 can be expressed in bacterial cells such as *C. glutamicum*, insect cells (using baculovirus expression vectors), yeast and other fungal cells (see Romanos, M.A. *et al.* (1992) "Foreign gene expression in yeast: a review", *Yeast* 8: 423-488; van den Hondel,

C.A.M.J.J. *et al.* (1991) "Heterologous gene expression in filamentous fungi" in: More Gene Manipulations in Fungi, J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego; and van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, Peberdy, J.F. *et al.*, eds., p. 1-28, Cambridge University Press: Cambridge), algae and multicellular plant cells (see Schmidt, R. and Willmitzer, L. (1988) High efficiency *Agrobacterium tumefaciens* -mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants" *Plant Cell Rep.*: 583-586), or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein but also to the C-terminus or fused within suitable regions in the proteins. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the HA protein is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-X protein. The fusion protein can be purified by affinity chromatography using glutathione-agarose resin.

Recombinant HA protein unfused to GST can be recovered by cleavage of the fusion protein with thrombin.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) pLG338, pACYC184, pBR322, pUC18, 5 pUC19, pKC30, pRep4, pHS1, pHS2, pPLc236, pMBL24, pLG200, pUR290, pIN-III113-B1,  $\lambda$ gt11, pBdCl, and pET 11d (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89 ; and Pouwels *et al.*, eds. (1985) *Cloning Vectors*. Elsevier: New York IBSN 0 444 904018). Target gene expression from the pTrc vector relies on host RNA polymerase 10 transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident  $\lambda$  prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter. For transformation 15 of other varieties of bacteria, appropriate vectors may be selected. For example, the plasmids pIJ101, pIJ364, pIJ702 and pIJ361 are known to be useful in transforming *Streptomyces*, while plasmids pUB110, pC194, or pBD214 are suited for transformation of *Bacillus* species. Several plasmids of use in the transfer of genetic information into *Corynebacterium* include pHM1519, pBL1, pSA77, or pAJ667 (Pouwels *et al.*, eds. 20 (1985) *Cloning Vectors*. Elsevier: New York IBSN 0 444 904018).

One strategy to maximize recombinant protein expression is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another 25 strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in the bacterium chosen for expression, such as *C. glutamicum* (Wada *et al.* (1992) *Nucleic Acids Res.* 20: 2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

30 In another embodiment, the HA protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYepSec1 (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), 2  $\mu$ , pAG-1, Yep6, Yep13, pEMBLYe23,

- pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and methods for the construction of vectors appropriate for use in other fungi, such as the filamentous fungi, include those detailed in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi, in: Applied Molecular Genetics of Fungi, J.F. Peberdy, *et al.*, eds., p. 1-28, Cambridge University Press: Cambridge, and Pouwels *et al.*, eds. (1985) *Cloning Vectors*. Elsevier: New York (IBSN 0 444 904018).

Alternatively, the HA proteins of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (*e.g.*, Sf 9 cells) include the pAc series (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In another embodiment, the HA proteins of the invention may be expressed in unicellular plant cells (such as algae) or in plant cells from higher plants (*e.g.*, the spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197; and Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acid. Res.* 12: 8711-8721, and include pLGV23, pGHIac+, pBIN19, pAK2004, and pDH51 (Pouwels *et al.*, eds. (1985) *Cloning Vectors*. Elsevier: New York IBSN 0 444 904018).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring

Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji *et al.* (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund *et al.* (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the  $\alpha$ -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to HA mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene

expression using antisense genes see Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews - Trends in Genetics*, Vol. 1(1) (1986).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic or eukaryotic cell. For example, an HA protein can be expressed in bacterial cells such as *C. glutamicum*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other suitable host cells are known to those of ordinary skill in the art. Microorganisms related to *Corynebacterium glutamicum* which may be conveniently used as host cells for the nucleic acid and protein molecules of the invention are set forth in Table 3.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection", "conjugation" and "transduction" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (*e.g.*, linear DNA or RNA (*e.g.*, a linearized vector or a gene construct alone without a vector) or nucleic acid in the form of a vector (*e.g.*, a plasmid, phage, phasmid, phagemid, transposon or other DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, natural competence, chemical-mediated transfer, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual. 2nd, ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (*e.g.*, resistance to antibiotics) is

generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding an HA protein or can be  
5 introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by, for example, drug selection (*e.g.*, cells that have incorporated the selectable marker gene will survive, while the other cells die).

To create a homologous recombinant microorganism, a vector is prepared which contains at least a portion of an HA gene into which a deletion, addition or substitution  
10 has been introduced to thereby alter, *e.g.*, functionally disrupt, the HA gene. Preferably, this HA gene is a *Corynebacterium glutamicum* HA gene, but it can be a homologue from a related bacterium or even from a mammalian, yeast, or insect source. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous HA gene is functionally disrupted (*i.e.*, no longer  
15 encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous HA gene is mutated or otherwise altered but still encodes functional protein (*e.g.*, the upstream regulatory region can be altered to thereby alter the expression of the endogenous HA protein). In the homologous recombination vector, the altered portion  
20 of the HA gene is flanked at its 5' and 3' ends by additional nucleic acid of the HA gene to allow for homologous recombination to occur between the exogenous HA gene carried by the vector and an endogenous HA gene in a microorganism. The additional flanking HA nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA  
25 (both at the 5' and 3' ends) are included in the vector (see *e.g.*, Thomas, K.R., and Capecchi, M.R. (1987) Cell 51: 503 for a description of homologous recombination vectors). The vector is introduced into a microorganism (*e.g.*, by electroporation) and cells in which the introduced HA gene has homologously recombined with the endogenous HA gene are selected, using art-known techniques.

30 In another embodiment, recombinant microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene. For example, inclusion of an HA gene on a vector placing it under control of the lac



- 52 -

operon permits expression of the HA gene only in the presence of IPTG. Such regulatory systems are well known in the art.

In another embodiment, an endogenous HA gene in a host cell is disrupted (*e.g.*, by homologous recombination or other genetic means known in the art) such that  
5 expression of its protein product does not occur. In another embodiment, an endogenous or introduced HA gene in a host cell has been altered by one or more point mutations, deletions, or inversions, but still encodes a functional HA protein. In still another embodiment, one or more of the regulatory regions (*e.g.*, a promoter, repressor, or inducer) of an HA gene in a microorganism has been altered (*e.g.*, by deletion,  
10 truncation, inversion, or point mutation) such that the expression of the HA gene is modulated. One of ordinary skill in the art will appreciate that host cells containing more than one of the described HA gene and protein modifications may be readily produced using the methods of the invention, and are meant to be included in the present invention.

15 A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (*i.e.*, express) an HA protein. Accordingly, the invention further provides methods for producing HA proteins using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an HA protein has been  
20 introduced, or into which genome has been introduced a gene encoding a wild-type or altered HA protein) in a suitable medium until HA protein is produced. In another embodiment, the method further comprises isolating HA proteins from the medium or the host cell.

### 25 C. Isolated HA Proteins

Another aspect of the invention pertains to isolated HA proteins, and biologically active portions thereof. An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized.  
30 The language "substantially free of cellular material" includes preparations of HA protein in which the protein is separated from cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language

"substantially free of cellular material" includes preparations of HA protein having less than about 30% (by dry weight) of non-HA protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-HA protein, still more preferably less than about 10% of non-HA protein, and most preferably less than about 5% non-HA protein. When the HA protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, *i.e.*, culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of HA protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of HA protein having less than about 30% (by dry weight) of chemical precursors or non-HA chemicals, more preferably less than about 20% chemical precursors or non-HA chemicals, still more preferably less than about 10% chemical precursors or non-HA chemicals, and most preferably less than about 5% chemical precursors or non-HA chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the HA protein is derived. Typically, such proteins are produced by recombinant expression of, for example, a *C. glutamicum* HA protein in a microorganism such as *C. glutamicum*.

An isolated HA protein or a portion thereof of the invention can participate in the repair or recombination of DNA, in the transposition of genetic material, in gene expression (*i.e.*, the processes of transcription or translation), in protein folding, or in protein secretion in *Corynebacterium glutamicum*, or has one or more of the activities set forth in Table 1. In preferred embodiments, the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) such that the protein or portion thereof maintains the ability to participate in the maintenance of homeostasis in *C. glutamicum*, or to perform a function involved in the adaptation of this microorganism to different environmental conditions. The portion of the protein is preferably a biologically active portion as described herein. In another

preferred embodiment, an HA protein of the invention has an amino acid sequence set forth as an even-numbered SEQ ID NO: of the Sequence Listing. In yet another preferred embodiment, the HA protein has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to a nucleotide sequence the invention (*e.g.*, a sequence of an odd-numbered SEQ ID NO: of the Sequence Listing). In still another preferred embodiment, the HA protein has an amino acid sequence which is encoded by a nucleotide sequence that is at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to one of the nucleic acid sequences of the invention, or a portion thereof. Ranges and identity values intermediate to the above-recited values, (*e.g.*, 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. The preferred HA proteins of the present invention also preferably possess at least one of the HA activities described herein. For example, a preferred HA protein of the present invention includes an amino acid sequence encoded by a nucleotide sequence which hybridizes, *e.g.*, hybridizes under stringent conditions, to a nucleotide sequence of the invention, and which can participate in the maintenance of homeostasis in *C. glutamicum*, or can perform a function involved in the adaptation of this microorganism to different environmental conditions, or which has one or more of the activities set forth in Table 1.

In other embodiments, the HA protein is substantially homologous to an amino acid sequence of the invention (*e.g.*, a sequence of an even-numbered SEQ ID NO: of the Sequence Listing) and retains the functional activity of the protein of one of the amino acid sequences of the invention yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the HA protein is a protein which comprises an amino acid sequence which is at least about 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, or 60%, preferably at least about 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%,

69%, or 70%, more preferably at least about 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, or 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, or 90%, or 91%, 92%, 93%, 94%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to an entire amino acid sequence of the invention and which  
5 has at least one of the HA activities described herein. Ranges and identity values intermediate to the above-recited values, (e.g., 70-90% identical or 80-95% identical) are also intended to be encompassed by the present invention. For example, ranges of identity values using a combination of any of the above values recited as upper and/or lower limits are intended to be included. In another embodiment, the invention pertains  
10 to a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of the invention.

Biologically active portions of an HA protein include peptides comprising amino acid sequences derived from the amino acid sequence of an HA protein, e.g., the amino acid sequence of an even-numbered SEQ ID NO: of the Sequence Listing, the amino  
15 acid sequence of a protein homologous to an HA protein, which include fewer amino acids than a full length HA protein or the full length protein which is homologous to an HA protein, and exhibit at least one activity of an HA protein. Typically, biologically active portions (peptides, e.g., peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with  
20 at least one activity of an HA protein. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically active portions of an HA protein include one or more selected domains/motifs or portions thereof having biological activity.

25 HA proteins are preferably produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as described above) and the HA protein is expressed in the host cell. The HA protein can then be isolated from the cells by an appropriate purification scheme using standard  
30 protein purification techniques. Alternative to recombinant expression, an HA protein, polypeptide, or peptide can be synthesized chemically using standard peptide synthesis techniques. Moreover, native HA protein can be isolated from cells (e.g., endothelial

- 56 -

cells), for example using an anti-HA antibody, which can be produced by standard techniques utilizing an HA protein or fragment thereof of this invention.

The invention also provides HA chimeric or fusion proteins. As used herein, an HA "chimeric protein" or "fusion protein" comprises an HA polypeptide operatively  
5 linked to a non-HA polypeptide. An "HA polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an HA protein, whereas a "non-HA polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the HA protein, *e.g.*, a protein which is different from the HA protein and which is derived from the same or a different organism. Within the  
10 fusion protein, the term "operatively linked" is intended to indicate that the HA polypeptide and the non-HA polypeptide are fused in-frame to each other. The non-HA polypeptide can be fused to the N-terminus or C-terminus of the HA polypeptide. For example, in one embodiment the fusion protein is a GST-HA fusion protein in which the HA sequences are fused to the C-terminus of the GST sequences. Such fusion proteins  
15 can facilitate the purification of recombinant HA proteins. In another embodiment, the fusion protein is an HA protein containing a heterologous signal sequence at its N-terminus. In certain host cells (*e.g.*, mammalian host cells), expression and/or secretion of an HA protein can be increased through use of a heterologous signal sequence.

Preferably, an HA chimeric or fusion protein of the invention is produced by  
20 standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid  
25 undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric  
30 gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (*e.g.*, a GST polypeptide). An HA-

encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the HA protein.

Homologues of the HA protein can be generated by mutagenesis, *e.g.*, discrete point mutation or truncation of the HA protein. As used herein, the term "homologue" refers to a variant form of the HA protein which acts as an agonist or antagonist of the activity of the HA protein. An agonist of the HA protein can retain substantially the same, or a subset, of the biological activities of the HA protein. An antagonist of the HA protein can inhibit one or more of the activities of the naturally occurring form of the HA protein, by, for example, competitively binding to a downstream or upstream member of a biochemical cascade which includes the HA protein, by binding to a target molecule with which the HA protein interacts, such that no functional interaction is possible, or by binding directly to the HA protein and inhibiting its normal activity.

In an alternative embodiment, homologues of the HA protein can be identified by screening combinatorial libraries of mutants, *e.g.*, truncation mutants, of the HA protein for HA protein agonist or antagonist activity. In one embodiment, a variegated library of HA variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of HA variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential HA sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (*e.g.*, for phage display) containing the set of HA sequences therein. There are a variety of methods which can be used to produce libraries of potential HA homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding the desired set of potential HA sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, *e.g.*, Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of the HA protein coding can be used to generate a variegated population of HA fragments for screening and subsequent

selection of homologues of an HA protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an HA coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to  
5 form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the HA protein.

10 Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of HA homologues. The most widely used techniques, which are amenable to high through-put  
15 analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the  
20 frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify HA homologues (Arkin and Yourvan (1992) *PNAS* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

In another embodiment, cell based assays can be exploited to analyze a variegated HA library, using methods well known in the art.

25

#### *D. Uses and Methods of the Invention*

The nucleic acid molecules, proteins, protein homologues, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification of *C. glutamicum* and related organisms; mapping of  
30 genomes of organisms related to *C. glutamicum*; identification and localization of *C. glutamicum* sequences of interest; evolutionary studies; determination of HA protein regions required for function; modulation of an HA protein activity; modulation of the

metabolism of one or more inorganic compounds; modulation of the modification or degradation of one or more aromatic or aliphatic compounds; modulation of cell wall synthesis or rearrangements; modulation of enzyme activity or proteolysis; and modulation of cellular production of a desired compound, such as a fine chemical.

5           The HA nucleic acid molecules of the invention have a variety of uses. First, they may be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof. Also, they may be used to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes; by probing the  
10   extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is nonpathogenic, it is related to pathogenic species, such as *Corynebacterium diphtheriae*. *Corynebacterium diphtheriae*  
15   is the causative agent of diphtheria, a rapidly developing, acute, febrile infection which involves both local and systemic pathology. In this disease, a local lesion develops in the upper respiratory tract and involves necrotic injury to epithelial cells; the bacilli secrete toxin which is disseminated through this lesion to distal susceptible tissues of the body. Degenerative changes brought about by the inhibition of protein synthesis in  
20   these tissues, which include heart, muscle, peripheral nerves, adrenals, kidneys, liver and spleen, result in the systemic pathology of the disease. Diphtheria continues to have high incidence in many parts of the world, including Africa, Asia, Eastern Europe and the independent states of the former Soviet Union. An ongoing epidemic of diphtheria in the latter two regions has resulted in at least 5,000 deaths since 1990.

25           In one embodiment, the invention provides a method of identifying the presence or activity of *Corynebacterium diphtheriae* in a subject. This method includes detection of one or more of the nucleic acid or amino acid sequences of the invention (e.g., the sequences set forth in as odd-numbered or even-numbered SEQ ID NOs, respectively, in the Sequence Listing) in a subject, thereby detecting the presence or activity of  
30   *Corynebacterium diphtheriae* in the subject. *C. glutamicum* and *C. diphtheriae* are related bacteria, and many of the nucleic acid and protein molecules in *C. glutamicum*



- 60 -

are homologous to *C. diphtheriae* nucleic acid and protein molecules, and can therefore be used to detect *C. diphtheriae* in a subject.

The nucleic acid and protein molecules of the invention may also serve as markers for specific regions of the genome. This has utility not only in the mapping of the genome, but also for functional studies of *C. glutamicum* proteins. For example, to identify the region of the genome to which a particular *C. glutamicum* DNA-binding protein binds, the *C. glutamicum* genome could be digested, and the fragments incubated with the DNA-binding protein. Those which bind the protein may be additionally probed with the nucleic acid molecules of the invention, preferably with readily detectable labels; binding of such a nucleic acid molecule to the genome fragment enables the localization of the fragment to the genome map of *C. glutamicum*, and, when performed multiple times with different enzymes, facilitates a rapid determination of the nucleic acid sequence to which the protein binds. Further, the nucleic acid molecules of the invention may be sufficiently homologous to the sequences of related species such that these nucleic acid molecules may serve as markers for the construction of a genomic map in related bacteria, such as *Brevibacterium lactofermentum*.

The HA nucleic acid molecules of the invention are also useful for evolutionary and protein structural studies. The processes involved in adaptation and the maintenance of homeostasis in which the molecules of the invention participate are utilized by a wide variety of species; by comparing the sequences of the nucleic acid molecules of the present invention to those encoding similar enzymes from other organisms, the evolutionary relatedness of the organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme. This type of determination is of value for protein engineering studies and may give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

Manipulation of the HA nucleic acid molecules of the invention may result in the production of HA proteins having functional differences from the wild-type HA proteins. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or activity.

- 61 -

The invention provides methods for screening molecules which modulate the activity of an HA protein, either by interacting with the protein itself or a substrate or binding partner of the HA protein, or by modulating the transcription or translation of an HA nucleic acid molecule of the invention. In such methods, a microorganism  
5 expressing one or more HA proteins of the invention is contacted with one or more test compounds, and the effect of each test compound on the activity or level of expression of the HA protein is assessed.

The modulation of activity or number of HA proteins involved in cell wall biosynthesis or rearrangements may impact the production, yield, and/or efficiency of  
10 production of one or more fine chemicals from *C. glutamicum* cells. For example, by altering the activity of these proteins, it may be possible to modulate the structure or thickness of the cell wall. The cell wall serves in large measure as a protective device against osmotic lysis and external sources of injury; by modifying the cell wall it may be possible to increase the ability of *C. glutamicum* to withstand the mechanical and shear  
15 force stresses encountered by this microorganism during large-scale fermentor culture. Further, each *C. glutamicum* cell is surrounded by a thick cell wall, and thus, a significant portion of the biomass present in large scale culture consists of cell wall. By increasing the rate at which the cell wall is synthesized or by activating cell wall synthesis (through genetic engineering of the HA cell wall proteins of the invention) it  
20 may be possible to improve the growth rate of the microorganism. Similarly, by decreasing the activity or number of proteins involved in the degradation of cell wall or by decreasing the repression of cell wall biosynthesis, an overall increase in cell wall production may be achieved. An increase in the number of viable *C. glutamicum* cells (as may be accomplished by any of the foregoing described protein alterations) should  
25 result in increased numbers of cells producing the desired fine chemical in large-scale fermentor culture, which should permit increased yields or efficiency of production of these compounds from the culture.

The modulation of activity or number of *C. glutamicum* HA proteins that participate in the modification or degradation of aromatic or aliphatic compounds may  
30 also have direct or indirect impacts on the production of one or more fine chemicals from these cells. Certain aromatic or aliphatic modification or degradation products are desirable fine chemicals (e.g., organic acids or modified aromatic and aliphatic

compounds); thus, by modifying the enzymes which perform these modifications (*e.g.*, hydroxylation, methylation, or isomerization) or degradation reactions, it may be possible to increase the yields of these desired compounds. Similarly, by decreasing the activity or number of proteins involved in pathways which further degrade the modified or breakdown products of the aforementioned reactions it may be possible to improve the yields of these fine chemicals from *C. glutamicum* cells in culture.

These aromatic and aliphatic modification and degradative enzymes are themselves fine chemicals. In purified form, these enzymes may be used to degrade aromatic and aliphatic compounds (*e.g.*, toxic chemicals such as petroleum products), either for the bioremediation of polluted sites, for the engineered decomposition of wastes, or for the large-scale and economically feasible production of desired modified aromatic or aliphatic compounds or their breakdown products, some of which may be conveniently used as carbon or energy sources for other fine chemical-producing compounds in culture (see, *e.g.*, Faber, K. (1995) *Biotransformations in Organic Chemistry*, Springer: Berlin and references therein; and Roberts, S.M., ed. (1992-1996) *Preparative Biotransformations*, Wiley: Chichester, and references therein). By genetically altering these proteins such that their regulation by other cellular mechanisms is lessened or abolished, it may be possible to increase the overall number or activity of these proteins, thereby improving not only the yield of these fine chemicals but also the activity of these harvested proteins.

The modification of these aromatic and aliphatic modifying and degradation enzymes may also have an indirect effect on the production of one or more fine chemical. Many aromatic and aliphatic compounds (such as those that may be encountered as impurities in culture media or as waste products from cellular metabolism) are toxic to cells; by modifying and/or degrading these compounds such that they may be readily removed or destroyed, cellular viability should be increased. Further, these enzymes may modify or degrade these compounds in such a manner that the resulting products may enter the normal carbon metabolism pathways of the cell, thus rendering the cell able to use these compounds as alternate carbon or energy sources. In large-scale culture situations, when there may be limiting amounts of optimal carbon sources, these enzymes provide a method by which cells may continue to grow and divide using aromatic or aliphatic compounds as nutrients. In either case, the

resulting increase in the number of *C. glutamicum* cells in the culture producing the desired fine chemical should in turn result in increased yields or efficiency of production of the fine chemical(s).

Modifications in activity or number of HA proteins involved in the metabolism of inorganic compounds may also directly or indirectly affect the production of one or more fine chemicals from *C. glutamicum* or related bacterial cultures. For example, many desirable fine chemicals, such as nucleic acids, amino acids, cofactors and vitamins (*e.g.*, thiamine, biotin, and lipoic acid) cannot be synthesized without inorganic molecules such as phosphorous, nitrate, sulfate, and iron. The inorganic metabolism proteins of the invention permit the cell to obtain these molecules from a variety of inorganic compounds and to divert them into various fine chemical biosynthetic pathways. Therefore, by increasing the activity or number of enzymes involved in the metabolism of these inorganic compounds, it may be possible to increase the supply of these possibly limiting inorganic molecules, thereby directly increasing the production or efficiency of production of various fine chemicals from *C. glutamicum* cells containing such altered proteins. Modification of the activity or number of inorganic metabolism enzymes of the invention may also render *C. glutamicum* able to better utilize limited inorganic compound supplies, or to utilize nonoptimal inorganic compounds to synthesize amino acids, vitamins, cofactors, or nucleic acids, all of which are necessary for continued growth and replication of the cell. By improving the viability of these cells in large-scale culture, the number of *C. glutamicum* cells producing one or more fine chemicals in the culture may also be increased, in turn increasing the yields or efficiency of production of one or more fine chemicals.

*C. glutamicum* enzymes for general processes are themselves desirable fine chemicals. The specific properties of enzymes (*i.e.*, regio- and stereospecificity, among others) make them useful catalysts for chemical reactions *in vitro*. Either whole *C. glutamicum* cells may be incubated with an appropriate substrate such that the desired product is produced by enzymes in the cell, or the desired enzymes may be overproduced and purified from *C. glutamicum* cultures (or those of a related bacterium) and subsequently utilized in *in vitro* reactions in an industrial setting (either in solution or immobilized on a suitable immobile phase). In either situation, the enzyme can either be a natural *C. glutamicum* protein, or it may be mutagenized to have an altered activity;

typical industrial uses for such enzymes include as catalysts in the chemical industry (e.g., for synthetic organic chemistry) as food additives, as feed components, for fruit processing, for leather preparation, in detergents, in analysis and medicine, and in the textile industry (see, e.g., Yamada, H. (1993) "Microbial reactions for the production of useful organic compounds," *Chimica* 47: 5-10; Roberts, S.M. (1998) Preparative biotransformations: the employment of enzymes and whole-cells in synthetic chemistry," *J. Chem. Soc. Perkin Trans. 1*: 157-169; Zaks, A. and Dodds, D.R. (1997) "Application of biocatalysis and biotransformations to the synthesis of pharmaceuticals," *DDT* 2: 513-531; Roberts, S.M. and Williamson, N.M. (1997) "The use of enzymes for the preparation of biologically active natural products and analogues in optically active form," *Curr. Organ. Chemistry* 1: 1-20; Faber, K. (1995) Biotransformations in Organic Chemistry, Springer: Berlin; Roberts, S.M., ed. (1992-96) Preparative Biotransformations, Wiley: Chichester; Cheetham, P.S.J. (1995) "The applications of enzymes in industry" in : Handbook of Enzyme Biotechnology, 3<sup>rd</sup> ed., Wiseman, A., ed., Ellis: Horwood, p. 419-552; and Ullmann's Encyclopedia of Industrial Chemistry (1987), vol. A9, Enzymes, p. 390-457). Thus, by increasing the activity or number of these enzymes, it may be possible to also increase the ability of the cell to convert supplied substrates to desired products, or to overproduce these enzymes for increased yields in large-scale culture. Further, by mutagenizing these proteins it may be possible to remove feedback inhibition or other repressive cellular regulatory controls such that greater numbers of these enzymes may be produced and activated by the cell, thereby leading to greater yields, production, or efficiency of production of these fine chemical proteins from large-scale cultures. Further, manipulation of these enzymes may alter the activity of one or more *C. glutamicum* metabolic pathways, such as those for the biosynthesis or secretion of one or more fine chemicals.

Mutagenesis of the proteolytic enzymes of the invention such that they are altered in activity or number may also directly or indirectly affect the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*. For example, by increasing the activity or number of these proteins, it may be possible to increase the ability of the bacterium to survive in large-scale culture, due to an increased ability of the cell to rapidly degrade proteins misfolded in response to the high temperatures, nonoptimal pH, and other stresses encountered during fermentor culture.

- Increased numbers of cells in these cultures may result in increased yields or efficiency of production of one or more desired fine chemicals, due to the relatively larger number of cells producing these compounds in the culture. Also, *C. glutamicum* cells possess multiple cell-surface proteases which serve to break down external nutrients into
- 5 molecules which may be more readily incorporated by the cells as carbon/energy sources or nutrients of other kinds. An increase in activity or number of these enzymes may improve this turnover and increase the levels of available nutrients, thereby improving cell growth or production. Thus, modifications of the proteases of the invention may indirectly impact *C. glutamicum* fine chemical production.
- 10 A more direct impact on fine chemical production in response to the modification of one or more of the proteases of the invention may occur when these proteases are involved in the production or degradation of a desired fine chemical. By decreasing the activity of a protease which degrades a fine chemical or a protein involved in the synthesis of a fine chemical it may be possible to increase the levels of
- 15 that fine chemical (due to the decreased degradation or increased synthesis of the compound). Similarly, by increasing the activity of a protease which degrades a compound to result in a fine chemical or a protein involved in the degradation of a fine chemical, a similar result should be achieved: increased levels of the desired fine chemical from *C. glutamicum* cells containing these engineered proteins.
- 20 The aforementioned mutagenesis strategies for HA proteins to result in increased yields of a fine chemical from *C. glutamicum* are not meant to be limiting; variations on these strategies will be readily apparent to one of ordinary skill in the art. Using such strategies, and incorporating the mechanisms disclosed herein, the nucleic acid and protein molecules of the invention may be utilized to generate *C. glutamicum* or related
- 25 strains of bacteria expressing mutated HA nucleic acid and protein molecules such that the yield, production, and/or efficiency of production of a desired compound is improved. This desired compound may be any product produced by *C. glutamicum*, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally
- 30 occur in the metabolism of *C. glutamicum*, but which are produced by a *C. glutamicum* strain of the invention.

- 66 -

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents, published patent applications, Tables, and the sequence listing cited throughout this application are hereby incorporated by reference.

Table 1: Genes in the Application

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
1	2	RXA02548	GR00727	3	293	SULFATE ADENYLATE TRANSFERASE SUBUNIT 2 (EC 2.7.7.4)
3	4	RXN00249	VW0057	36825	35869	ADENYL SULFATE KINASE (EC 2.7.1.25)
5	6	F RXA00249	GR0037	8837	7884	ADENYL SULFATE KINASE (EC 2.7.1.25)
7	8	RXA01073	GR00300	1274	2104	NH(3)-DEPENDENT NAD(+) SYNTHETASE (EC 6.3.5.1)

  

Urease						
Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
9	10	RXN02913	VW0020	8998	8513	UREASE BETA SUBUNIT (EC 3.5.1.5)
11	12	F RXA02264	GR00655	123	4	UREASE ALPHA SUBUNIT (EC 3.5.1.5)
13	14	RXN02274	VW0020	8509	6800	UREASE ALPHA SUBUNIT (EC 3.5.1.5)
15	16	F RXA02274	GR00656	3	1604	UREASE ALPHA SUBUNIT (EC 3.5.1.5)
17	18	RXA02265	GR00655	452	153	UREASE GAMMA SUBUNIT (EC 3.5.1.5)
19	20	RXA02278	GR00656	3420	4268	UREASE OPERON URED PROTEIN
21	22	GR00656	GR00656	1632	2102	UREASE ACCESSORY PROTEIN UREE
23	24	RXA02276	GR00656	2105	2782	UREASE ACCESSORY PROTEIN UREF
25	26	RXA02277	GR00656	2802	3416	UREASE ACCESSORY PROTEIN UREG
27	28	RXA02803	GR00742	7742	8737	4-HYDROXYBENZOATE OCTAPRENYLTRANSFERASE (EC 2.5.1.-)
29	30	RXA01385	GR00406	5320	3440	PHENOL 2 MONOOXYGENASE (EC 1.14.13.7)

  

Proteolysis						
Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
31	32	RXN00675	VW0005	33258	34049	METHIONINE AMINOPEPTIDASE (EC 3.4.11.18)
33	34	F RXA00675	GR00178	2	484	METHIONINE AMINOPEPTIDASE (EC 3.4.11.18)
35	36	RXA01609	GR00449	2740	3612	METHIONINE AMINOPEPTIDASE (EC 3.4.11.18)
37	38	RXA01358	GR00393	5337	6857	ATP-DEPENDENT PROTEASE LA (EC 3.4.21.53)
39	40	RXA01458	GR00420	3225	2176	ATP-DEPENDENT PROTEASE LA (EC 3.4.21.53)
41	42	RXA01654	GR00459	986	1981	(AL022121) putative alkaline serine protease [Mycobacterium tuberculosis]
43	44	RXN01868	VW0127	9980	11905	ZINC METALLOPROTEASE (EC 3.4.24.-)
45	46	F RXA01868	GR00534	1640	30	ZINC METALLOPROTEASE (EC 3.4.24.-)
47	48	F RXA01869	GR00534	1954	1652	ZINC METALLOPROTEASE (EC 3.4.24.-)
49	50	RXN03028	VW0008	41156	43930	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPA
51	52	F RXA02470	GR00715	2216	3196	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPA



Table 1 (continued)

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
53	54	F RXA02471	GR00715	3159	4991	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPA
55	56	RXA02630	GR00748	2654	1332	(AL021999) putative serine protease [Mycobacterium tuberculosis]
57	58	RXA02834	GR00823	3	497	ATPases with chaperone activity, ATP-dependent protease subunit
59	60	RXA00112	GR00016	3687	2497	PROBABLE PERIPLASMIC SERINE PROTEASE DO-LIKE PRECURSOR
61	62	RXA00566	GR00152	742	137	ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT (EC 3.4.21.92)
63	64	RXA00567	GR00152	1388	798	ATP-DEPENDENT CLP PROTEASE PROTEOLYTIC SUBUNIT (EC 3.4.21.92)
65	66	RXN03094	VW0057	1794	43	CLPB PROTEIN
67	68	F RXA01668	GR00464	2205	3920	CLPB PROTEIN
69	70	RXN01120	VW0182	5678	4401	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX
71	72	F RXA01120	GR00310	2349	1072	ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX
73	74	RXA00744	GR00202	10722	9781	Periplasmic serine proteases
75	76	RXA00844	GR00228	3620	4453	Hypothetical Secretory Serine Protease (EC 3.4.21.-)
77	78	RXA01151	GR00324	862	5	ATP-dependent Zn proteases
79	80	RXA02317	GR00665	9664	9053	PEPTIDASE E (EC 3.4.-.-)
81	82	RXA02644	GR00751	767	117	XAA-PRO DIPEPTIDASE (EC 3.4.13.9)
83	84	RXN02820	VW0131	4799	6109	GAMMA-GLUTAMYL TRANSPEPTIDASE (EC 2.3.2.2)
85	86	F RXA02820	GR00801	1	507	GAMMA-GLUTAMYL TRANSPEPTIDASE (EC 2.3.2.2)
87	88	F RXA02000	GR00589	3430	3933	GAMMA-GLUTAMYL TRANSPEPTIDASE (EC 2.3.2.2)
89	90	RXN03176	VW0334	921	121	PENICILLIN-BINDING PROTEIN 5* PRECURSOR (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE) (EC 3.4.16.4)
91	92	F RXA02859	GR10005	846	121	PENICILLIN-BINDING PROTEIN 5* PRECURSOR (D-ALANYL-D-ALANINE CARBOXYPEPTIDASE) (EC 3.4.16.4)
93	94	RXA00137	GR00022	738	1826	XAA-PRO AMINOPEPTIDASE (EC 3.4.11.9)
95	96	RXN00499	VW0086	8158	9438	PROLINE IMINOPEPTIDASE (EC 3.4.11.5)
97	98	F RXA00499	GR00125	3	959	PROLINE IMINOPEPTIDASE
99	100	RXN00877	VW0099	2221	3885	PEPTIDYL-DIPEPTIDASE DCP (EC 3.4.15.5)
101	102	F RXA00877	GR00242	3	1067	PEPTIDYL-DIPEPTIDASE DCP (EC 3.4.15.5)
103	104	RXN01014	VW0209	13328	10728	AMINOPEPTIDASE N (EC 3.4.11.2)
105	106	F RXA01014	GR00289	3	1580	AMINOPEPTIDASE N (EC 3.4.11.2)
107	108	F RXA01018	GR00290	2289	3152	AMINOPEPTIDASE N (EC 3.4.11.2)
109	110	RXA01147	GR00323	1353	94	VACUOLAR AMINOPEPTIDASE I PRECURSOR (EC 3.4.11.1)
111	112	RXA01161	GR00329	1253	117	XAA-PRO AMINOPEPTIDASE (EC 3.4.11.9)
113	114	RXN01181	VW0065	1	957	AMINOPEPTIDASE A/I (EC 3.4.11.1)
115	116	F RXA01181	GR00337	1	957	AMINOPEPTIDASE
117	118	RXN01277	VW0009	32155	34158	PROLYL ENDOPEPTIDASE (EC 3.4.21.26)
119	120	F RXA01277	GR00368	1738	50	PROLYL ENDOPEPTIDASE (EC 3.4.21.26)
121	122	RXA01914	GR00548	125	550	AMINOPEPTIDASE
123	124	RXA02048	GR00624	207	1580	AMINOPEPTIDASE N (EC 3.4.11.2)
125	126	RXN00621	VW0135	5853	5071	PROTEASE II (EC 3.4.21.83)
127	128	F RXA00621	GR00163	4075	4857	PTRB periplasmic protease
129	130	RXN00622	VW0135	5150	3735	PROTEASE II (EC 3.4.21.83)
131	132	F RXA00622	GR00163	4778	6193	PTRB periplasmic protease
133	134	RXN00982	VW0149	7596	6091	(L42758) proteinase [Streptomyces lividans]
135	136	F RXA00977	GR00275	1647	2860	(L42758) proteinase [Streptomyces lividans]
137	138	F RXA00982	GR00276	5194	4949	(L42758) proteinase [Streptomyces lividans]
139	140	RXA00152	GR00023	7175	5880	HFLC PROTEIN (EC 3.4.-.-)
141	142	RXA02558	GR00731	4939	3965	HFLC PROTEIN (EC 3.4.-.-)

Table 1 (continued)

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
143	144	RXA00500	GR00125	969	1643	O-SIALOGLYCOPROTEIN ENDOPEPTIDASE (EC 3.4.24.57)
145	146	RXA00501	GR00125	1643	2149	O-SIALOGLYCOPROTEIN ENDOPEPTIDASE (EC 3.4.24.57)
147	148	RXA00502	GR00125	2156	3187	O-SIALOGLYCOPROTEIN ENDOPEPTIDASE (EC 3.4.24.57)

  

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
149	150	RXN02589	VV0098	16346	17110	Hypothetical Methyltransferase (EC 2.1.1.-)
151	152	F RXA02589	GR00741	13804	13040	Predicted S-adenosylmethionine-dependent methyltransferase
153	154	RXA00226	GR00032	26836	26012	SAM-dependent methyltransferases
155	156	RXN01885	VV0184	2004	2804	Hypothetical Methyltransferase (EC 2.1.1.-)
157	158	F RXA01885	GR00539	1589	2389	SAM-dependent methyltransferases
159	160	RXA02592	GR00741	18477	17707	SAM-dependent methyltransferases
161	162	RXN01795	VV0093	722	1318	MODIFICATION METHYLASE (EC 2.1.1.73)
163	164	F RXA01795	GR00507	706	1140	MODIFICATION METHYLASE (EC 2.1.1.73)
165	166	RXA01214	GR00351	1640	3130	LACCASE 1 PRECURSOR (EC 1.10.3.2)
167	168	RXA01250	GR00364	592	5	LACCASE 1 PRECURSOR (EC 1.10.3.2)
169	170	RXA02477	GR00715	10581	11201	CARBONIC ANHYDRASE (EC 4.2.1.1)
171	172	RXN00833	GR00225	374	6	THIOL PEROXIDASE (EC 1.11.1.-)
173	174	F RXA00833	GR00225	374	6	THIOL PEROXIDASE (EC 1.11.1.-)
175	176	RXA01224	GR00354	4186	5208	2-NITROPROPANE DIOXYGENASE (EC 1.13.11.32)
177	178	RXA01182	GR00337	1363	971	Hypothetical Oxidoreductase
179	180	RXA02531	GR00726	1226	1936	Hypothetical Oxidoreductase
181	182	RXN00689	VV0005	22416	20926	BETAINE-ALDEHYDE DEHYDROGENASE PRECURSOR (EC 1.2.1.8)
183	184	F RXA00689	GR00180	1401	775	BETAINE-ALDEHYDE DEHYDROGENASE PRECURSOR (EC 1.2.1.8)
185	186	RXN03128	VV0120	3	857	MORPHINE 6-DEHYDROGENASE (EC 1.1.1.218)
187	188	F RXA02192	GR00643	2	523	MORPHINE 6-DEHYDROGENASE (EC 1.1.1.218)
189	190	RXA02351	GR00679	132	1070	NITRILOTRIACETATE MONOOXYGENASE COMPONENT A (EC 1.14.13.-)
191	192	RXN00905	VV0238	8075	8875	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
193	194	F RXA00905	GR00247	2	694	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
195	196	RXA00906	GR00247	630	1133	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
197	198	RXA00907	GR00247	1143	1265	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
199	200	RXA02101	GR00631	3104	1842	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
201	202	RXN02565	VV0154	14299	13034	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
203	204	F RXA02565	GR00733	1	342	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
205	206	F RXA02567	GR00734	3	740	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
207	208	RXN03077	VV0043	1729	2913	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
209	210	F RXA02855	GR10002	1693	2877	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
211	212	RXA00026	GR00003	3657	5042	Hypothetical Amidohydrolase (EC 3.5.1.-)
213	214	RXA01971	GR00569	963	133	Hypothetical Metal-Dependent Hydrolase
215	216	RXA01802	GR00509	3461	4291	Predicted hydrolases (HAD superfamily)
217	218	RXN00866	VV0258	3557	4522	Predicted Zn-dependent hydrolases
219	220	F RXA00866	GR00236	3555	4499	Predicted Zn-dependent hydrolases
221	222	RXA02410	GR00703	792	127	Predicted Zn-dependent hydrolases

## Enzymes in general

Table 1 (continued)

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
223	224	RXA00961	GR00267	2	433	SALICYLATE HYDROXYLASE (EC 1.14.13.1)
225	226	RXA00111	GR00016	930	1922	SOLUBLE EPOXIDE HYDROLASE (SEH) (EC 3.3.2.3)
227	228	RXA01932	GR00555	6479	5583	ACETYL-HYDROLASE (EC 3.1.1.-)
229	230	RXA02574	GR00739	833	1840	PUTATIVE SECRETED HYDROLASE
231	232	RXA00983	WV0231	1796	321	SIALIDASE PRECURSOR (EC 3.2.1.18)
233	234	F RXA00983	GR00278	1200	4	SIALIDASE PRECURSOR (EC 3.2.1.18)
235	236	RXA00984	GR00278	1716	1300	SIALIDASE PRECURSOR (EC 3.2.1.18)
237	238	RXA02513	WV0193	737	6	SIALIDASE PRECURSOR (EC 3.2.1.18)
239	240	F RXA02513	GR00722	93	824	SIALIDASE PRECURSOR (EC 3.2.1.18)
241	242	RXA00903	GR00246	637	5	Putative epimerase
243	244	RXA01224	GR00354	4186	5208	2-NITROPROPANE DIOXYGENASE (EC 1.13.11.32)
245	246	RXA01571	GR00438	1360	1959	ALCOHOL DEHYDROGENASE (EC 1.1.1.1)
247	248	RXA02478	WV0119	7564	6350	SIALIDASE PRECURSOR (EC 3.2.1.18)
249	250	RXA00343	WV0125	1118	6	3-OXOSTEROID 1-DEHYDROGENASE (EC 1.3.99.4)
251	252	RXA01555	WV0135	29820	28861	3-OXOSTEROID 1-DEHYDROGENASE (EC 1.3.99.4)
253	254	RXA01166	WV0117	18142	16838	EXTRACELLULAR LIPASE PRECURSOR (EC 3.1.1.3)
255	256	RXA02001	WV0326	630	1787	N-ACYL-L-AMINO ACID AMIDOHYDROLASE (EC 3.5.1.14)
257	258	RXA03145	WV0142	7561	7115	4-OXALOCROTONATE TAUTOMERASE (EC 5.3.2.-)
259	260	RXA01466	WV0019	7050	6091	ARYLESTERASE (EC 3.1.1.2)
261	262	RXA01145	WV0077	7538	6525	KETOL-ACID REDUCTOISOMERASE (EC 1.1.1.86)
263	264	RXA03088	WV0052	3431	3817	Hypothetical Methyltransferase (EC 2.1.1.-)
265	266	RXA02952	WV0320	1032	1547	PUTATIVE REDUCTASE
267	268	RXA00513	WV0092	1573	653	CARBOXYVINYL-CARBOXYPHOSPHONATE PHOSPHORYLMUTASE (EC 2.7.8.23)
269	270	RXA01152	WV0136	1740	907	PROTEIN-L-ISOASPARTATE O-METHYLTRANSFERASE (EC 2.1.1.77)
271	272	RXA00787	WV0321	3736	5637	D-AMINO ACID DEHYDROGENASE LARGE SUBUNIT (EC 1.4.99.1)

## N-metabolism

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
273	274	RXA01302	WV0148	2837	2385	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
275	276	F RXA01302	GR00376	370	5	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
277	278	RXA01308	WV0148	2406	4	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
279	280	F RXA01307	GR00377	686	6	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
281	282	F RXA01308	GR00378	1211	6	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
283	284	RXA01309	WV0158	1	801	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
285	286	F RXA01309	GR00379	719	51	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
287	288	RXA02017	GR00610	1731	1048	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
289	290	RXA02018	GR00610	2788	1739	NITRATE REDUCTASE ALPHA CHAIN (EC 1.7.99.4)
291	292	RXA02016	GR00610	1036	260	NITRATE REDUCTASE BETA CHAIN (EC 1.7.99.4)
293	294	RXA00471	GR00119	2997	3686	NITRATE REDUCTASE GAMMA CHAIN (EC 1.7.99.4)
295	296	RXA00133	GR00021	201	1013	NITRATE/NITRITE RESPONSE REGULATOR PROTEIN NARL
297	298	RXA00650	GR00169	4017	3382	NITRATE/NITRITE RESPONSE REGULATOR PROTEIN NARP

Table 1 (continued)

<u>Nucleic Acid SEQ ID NO</u>	<u>Amino Acid SEQ ID NO</u>	<u>Identification Code</u>	<u>Contig.</u>	<u>NT Start</u>	<u>NT Stop</u>	<u>Function</u>
299	300	RXA01189	GR00339	2545	1937	NITRATE/NITRITE RESPONSE REGULATOR PROTEIN NARP
301	302	RXA01607	GR00449	123	752	NITRATE/NITRITE RESPONSE REGULATOR PROTEIN NARP
303	304	RXN00470	VV0086	27401	28669	NITRATE/NITRITE SENSOR PROTEIN NARX (EC 2.7.3.-)
305	306	F RXA00470	GR00119	1752	2951	NITRATE/NITRITE SENSOR PROTEIN NARX (EC 2.7.3.-)
307	308	RXA00756	GR00203	2932	1937	N UTILIZATION SUBSTANCE PROTEIN A
309	310	RXA00139	GR00022	2514	3224	N UTILIZATION SUBSTANCE PROTEIN B
311	312	RXA01303	GR00376	1724	390	NITRITE EXTRUSION PROTEIN
313	314	RXA01412	GR00412	620	417	NITROGEN FIXATION PROTEIN FIXI (PROBABLE E1-E2 TYPE CATION ATPASE) (EC 3.6.1.-)
315	316	RXA00773	GR00205	3208	4350	NITROGEN REGULATION PROTEIN NIFR3
317	318	RXA02746	GR00764	1	267	NITROGEN REGULATORY PROTEIN P-II
319	320	RXA02745	GR00763	15350	14472	MODULATION ATP-BINDING PROTEIN I
321	322	RXN00820	VV0054	19455	19817	MODULATION PROTEIN N
323	324	F RXA00820	GR00221	1007	1369	MODULATION PROTEIN N
325	326	RXA01059	GR00296	8782	9390	OXYGEN-INSENSITIVE NAD(P)H NITROREDUCTASE (EC 1.-.-.-)
327	328	RXN01386	VV0008	39246	38317	NITRILASE REGULATOR
329	330	RXN00073	VV0154	2369	687	FERREDOXIN-NITRITE REDUCTASE (EC 1.7.7.1)
331	332	RXN03131	VV0127	276	4	RHIZOPINE CATABOLISM PROTEIN MOCC
333	334	RXS00153	VV0167	4195	4620	MODULATION PROTEIN

**Urease**

<u>Nucleic Acid SEQ ID NO</u>	<u>Amino Acid SEQ ID NO</u>	<u>Identification Code</u>	<u>Contig.</u>	<u>NT Start</u>	<u>NT Stop</u>	<u>Function</u>
-----------------------------------	---------------------------------	----------------------------	----------------	-----------------	----------------	-----------------

**Phosphate and Phosphonate metabolism**

<u>Nucleic Acid SEQ ID NO</u>	<u>Amino Acid SEQ ID NO</u>	<u>Identification Code</u>	<u>Contig.</u>	<u>NT Start</u>	<u>NT Stop</u>	<u>Function</u>
335	336	RXN01716	VV0319	3259	2774	EXOPOLYPHOSPHATASE (EC 3.6.1.11)
337	338	RXN02972	VV0319	2763	2353	EXOPOLYPHOSPHATASE (EC 3.6.1.11)
339	340	RXN00663	VV0142	10120	11493	PHOH PROTEIN HOMOLOG
341	342	RXN00778	VV0103	18126	19250	PHOSPHATE-BINDING PERIPLASMIC PROTEIN PRECURSOR
343	344	RXN00250	VV0189	286	1032	DEDA PROTEIN - ALKALINE PHOSPHATASE LIKE PROTEIN

**Sulfate metabolism**

<u>Nucleic Acid SEQ ID NO</u>	<u>Amino Acid SEQ ID NO</u>	<u>Identification Code</u>	<u>Contig.</u>	<u>NT Start</u>	<u>NT Stop</u>	<u>Function</u>
345	346	RXA00072	GR00012	446	6	PHOSPHADENOSINE PHOSPHOSULFATE REDUCTASE (EC 1.8.99.4)
347	348	RXA00793	GR00211	1489	2644	SULFATE STARVATION-INDUCED PROTEIN 6
349	350	RXA01192	GR00342	161	733	SULFATE STARVATION-INDUCED PROTEIN 6

Table 1 (continued)

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
351	352	RXA00715	GR00188	2120	2914	THIOSULFATE SULFURTRANSFERASE (EC 2.8.1.1)
353	354	RXA01664	GR00463	1306	485	THIOSULFATE SULFURTRANSFERASE (EC 2.8.1.1)
355	356	RXN02334	VW0141	7939	7217	THIOSULFATE SULFURTRANSFERASE (EC 2.8.1.1)
357	358	F RXA02334	GR00672	2	355	THIOSULFATE SULFURTRANSFERASE (EC 2.8.1.1)

  

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
359	360	RXN01499	VW0008	7034	3213	ENTEROBACTIN SYNTHETASE COMPONENT F
361	362	RXN01997	VW0084	33308	33793	FERRITIN

  

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
363	364	RXA01848	GR00524	1532	789	MAGNESIUM-CHELATASE SUBUNIT CHLI
365	366	RXN01849	VW0139	16415	17515	MAGNESIUM-CHELATASE SUBUNIT CHLI
367	368	F RXA01849	GR00524	2004	1555	MAGNESIUM-CHELATASE SUBUNIT CHLI
369	370	F RXA01891	GR00474	570	4	MAGNESIUM-CHELATASE SUBUNIT CHLI
371	372	RXN00665	VW0252	135	635	MG2+/CITRATE COMPLEX SECONDARY TRANSPORTER

  

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
373	374	RXN03026	VW0007	28635	28901	3-DEHYDROQUINATE DEHYDRATASE (EC 4.2.1.10)
375	376	RXN02908	VW0025	8507	8247	O-SUCCINYLBENZOIC ACID--COA LIGASE (EC 6.2.1.26)
377	378	RXN03000	VW0235	570	4	SALICYLATE HYDROXYLASE (EC 1.14.13.1)
379	380	RXN03036	VW0014	671	6	PROTocatechuate 3,4-dioxygenase beta chain (EC 1.13.11.3)
381	382	RXN02974	VW0229	12631	12437	4-NITROPHENYLPHOSPHATASE (EC 3.1.3.41)
383	384	RXN00393	VW0025	7241	6348	1,4-DIHYDROXY-2-NAPHTHOATE OCTAPRENYLTRANSFERASE (EC 2.5.1.1)
385	386	RXN00948	VW0107	4266	5384	12-oxophylodienoate reductase (EC 1.3.1.42)
387	388	RXN01923	VW0020	3384	4133	2-HYDROXY-6-OXO-6-PHENYLHEXA-2,4-DIENOATE HYDROLASE (EC 3.7.1.1)
389	390	RXN00398	VW0025	14633	13884	2-PYRONE-4,6-DICARBOXYLATE LACTONASE (EC 3.1.1.57)
391	392	RXN02813	VW0128	13120	14118	3-CARBOXY-CIS,CIS-MUCONATE CYCLOISOMERASE HOMOLOG (EC 5.5.1.2)
393	394	RXN00136	VW0134	13373	14467	3-DEHYDROQUINATE SYNTHASE (EC 4.6.1.3)
395	396	RXN02508	VW0007	26733	28586	3-DEHYDROSHIKIMATE DEHYDRATASE (EC 4.2.1.1)
397	398	RXN02839	VW0322	3	449	4-HYDROXYBENZOATE OCTAPRENYLTRANSFERASE (EC 2.5.1.1)
399	400	RXN00639	VW0128	7858	8712	CATECHOL 1,2-DIOXYGENASE (EC 1.13.11.1)
401	402	RXN02530	VW0057	5469	6125	DIMETHYLANILINE MONOOXYGENASE (N-OXIDE FORMING) 1 (EC 1.14.13.8)
403	404	RXN00434	VW0112	12078	11212	QUINONE OXIDOREDUCTASE (EC 1.6.5.5)
405	406	RXN01619	VW0050	24649	23675	QUINONE OXIDOREDUCTASE (EC 1.6.5.5)
407	408	RXN01842	VW0234	1615	2532	QUINONE OXIDOREDUCTASE (EC 1.6.5.5)

## Modification and degradation of aromatic compounds

Table 1 (continued)

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
409	410	RXN00641	VV0128	7440	5950	TOLUATE 1,2-DIOXYGENASE ALPHA SUBUNIT (EC 1.14.12.-)
411	412	RXN01993	VV0182	16	1143	VANILLATE DEMETHYLASE (EC 1.14.-)
413	414	RXN00658	VV0083	15705	16397	PHENOL 2-MONOXYGENASE (EC 1.14.13.7)
415	416	RXN00178	VV0174	14670	15584	hydroxyquinol 1,2-dioxygenase (EC 1.13.11.37)
417	418	RXN01461	VV0128	12414	13025	PROTocatechuate 3,4-dioxygenase ALPHA CHAIN (EC 1.13.11.3)
419	420	RXN01653	VV0321	12867	11407	DIBENZOTHIOPHENE DESULFURIZATION ENZYME A
421	422	RXN02053	VV0009	39448	40026	DRGA PROTEIN
423	424	RXN00177	VV0174	13589	14656	MALEYLACETATE REDUCTASE (EC 1.3.1.32)
425	426	RXC00963	VV0249	1816	2652	PROTEIN involved in degradation of aromatic compounds

### Modification and degradation of aliphatic compounds

Nucleic Acid SEQ ID NO	Amino Acid SEQ ID NO	Identification Code	Contig.	NT Start	NT Stop	Function
427	428	RXN00299	VV0176	43379	42402	ALKANAL MONOOXYGENASE ALPHA CHAIN (EC 1.14.14.3)
429	430	F RXA00299	GR00048	7376	6633	ALKANAL MONOOXYGENASE ALPHA CHAIN (EC 1.14.14.3)
431	432	RXA00332	GR00057	16086	15385	ALKANAL MONOOXYGENASE ALPHA CHAIN (EC 1.14.14.3)
433	434	RXA01838	GR00519	2	820	ALKANAL MONOOXYGENASE ALPHA CHAIN (EC 1.14.14.3)
435	436	RXA02643	GR00750	1603	560	ALKANAL MONOOXYGENASE ALPHA CHAIN (EC 1.14.14.3)
437	438	RXA01933	GR00555	6590	7192	2-HALOALKANOIC ACID DEHALOGENASE I (EC 3.8.1.2)
439	440	RXA02351	GR00679	132	1070	NITRILOTRIACETATE MONOOXYGENASE COMPONENT A (EC 1.14.13.-)

TABLE 2 - Excluded Genes

GenBank™ Accession No.	Gene Name	Gene Function	Reference
A09073	ppg	Phosphoenol pyruvate carboxylase	Bachmann, B. et al. "DNA fragment coding for phosphoenolpyruvate carboxylase, recombinant DNA carrying said fragment, strains carrying the recombinant DNA and method for producing L-amino acids using said strains," Patent: EP 0358940-A 3 03/21/90
A45579, A45581, A45583, A45585 A45587		Threonine dehydratase	Moockel, B. et al. "Production of L-isoleucine by means of recombinant micro-organisms with deregulated threonine dehydratase," Patent: WO 9519442-A 5 07/20/95
AB003132	murC; ftsQ; ftsZ		Kobayashi, M. et al. "Cloning, sequencing, and characterization of the ftsZ gene from coryneform bacteria," <i>Biochem. Biophys. Res. Commun.</i> , 236(2):383-388 (1997)
AB015023	murC; ftsQ		Wachi, M. et al. "A murC gene from Coryneform bacteria," <i>Appl. Microbiol. Biotechnol.</i> , 51(2):223-228 (1999)
AB018530	ftsR		Kimura, E. et al. "Molecular cloning of a novel gene, ftsR, which rescues the detergent sensitivity of a mutant derived from <i>Brevibacterium lactofermentum</i> ," <i>Biosci. Biotechnol. Biochem.</i> , 60(10):1565-1570 (1996)
AB018531	ftsR1; ftsR2		
AB020624	murI	D-glutamate racemase	
AB023377	tkt	transketolase	
AB024708	glfB; glfD	Glutamine 2-oxoglutarate aminotransferase large and small subunits	
AB025424	acn	aconitase	
AB027714	rep	Replication protein	
AB027715	rep; aad	Replication protein; aminoglycoside adenyltransferase	
AF005242	argC	N-acetylglutamate-5-semialdehyde dehydrogenase	
AF005635	glnA	Glutamine synthetase	
AF030405	hisF	cyclase	
AF030520	argG	Argininosuccinate synthetase	
AF031518	argF	Ornithine carbamoyltransferase	
AF036932	aroD	3-dehydroquinate dehydratase	
AF038548	pyc	Pyruvate carboxylase	

Table 2 (continued)

	dcfAE; apt; rel	Dipeptide-binding protein; adenine phosphoribosyltransferase; GTP pyrophosphokinase	Wehmeier, L. et al. "The role of the Corynebacterium glutamicum rel gene in (p)ppGpp metabolism," <i>Microbiology</i> , 144:1853-1862 (1998)
AF038651			
AF041436	argR	Arginine repressor	
AF045998	impA	Inositol monophosphate phosphatase	
AF048764	argH	Argininosuccinate lyase	
AF049897	argC; argI; argB; argD; argF; argR; argG; argH	N-acetylglutamylphosphate reductase; ornithine acetyltransferase; N- acetylglutamate kinase; acetylornithine transaminase; ornithine carbamoyltransferase; arginine repressor; argininosuccinate synthase; argininosuccinate lyase	
AF050109	inhA	Enoyl-acyl carrier protein reductase	
AF050166	hisG	ATP phosphoribosyltransferase	
AF051846	hisA	Phosphoribosylformimino-5-amino-1- phosphoribosyl-4-imidazolecarboxamide isomerase	
AF052652	metA	Homoserine O-acetyltransferase	Park, S. et al. "Isolation and analysis of metA, a methionine biosynthetic gene encoding homoserine acetyltransferase in Corynebacterium glutamicum," <i>Mol. Cells</i> , 8(3):286-294 (1998)
AF053071	aroB	Dehydroquinase synthetase	
AF060558	hisH	Glutamine amidotransferase	
AF086704	hisE	Phosphoribosyl-ATP- pyrophosphohydrolase	
AF114233	aroA	5-enolpyruvylshikimate 3-phosphate synthase	
AF116184	panD	L-aspartate-alpha-decarboxylase precursor	Dusch, N. et al. "Expression of the Corynebacterium glutamicum panD gene encoding L-aspartate-alpha-decarboxylase leads to pantothenate overproduction in Escherichia coli," <i>Appl. Environ. Microbiol.</i> , 65(4):1530- 1539 (1999)
AF124518	aroD; aroE	3-dehydroquinase; shikimate dehydrogenase	
AF124600	aroC; aroK; aroB; pepQ	Chorismate synthase; shikimate kinase; 3- dehydroquinase synthase; putative cytoplasmic peptidase	
AF145897	inhA		
AF145898	inhA		



Table 2 (continued)

AJ001436	ectP	Transport of ectoine, glycine betaine, proline	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes: Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998)
AJ004934	dapD	Tetrahydrodipicolinate succinylase (incomplete)	Wehrmann, A. et al. "Different modes of diaminopimelate synthesis and their role in cell wall integrity: A study with <i>Corynebacterium glutamicum</i> ," <i>J. Bacteriol.</i> , 180(12):3159-3165 (1998)
AJ007732	ppc; secG; amt; ocd; soxA	Phosphoenolpyruvate-carboxylase; ?; high affinity ammonium uptake protein; putative ornithine-cyclodecarboxylase; sarcosine oxidase	
AJ010319	ftsY, glnB, glnD; srp; amtP	Involved in cell division; PII protein; uridylyltransferase (uridylyl-removing enzyme); signal recognition particle; low affinity ammonium uptake protein	Jakoby, M. et al. "Nitrogen regulation in <i>Corynebacterium glutamicum</i> ; Isolation of genes involved in biochemical characterization of corresponding proteins," <i>FEMS Microbiol.</i> , 173(2):303-310 (1999)
AJ132968	cat	Chloramphenicol acetyl transferase	
AJ224946	mgo	L-malate: quinone oxidoreductase	Molenaar, D. et al. "Biochemical and genetic characterization of the membrane-associated malate dehydrogenase (acceptor) from <i>Corynebacterium glutamicum</i> ," <i>Eur. J. Biochem.</i> , 254(2):395-403 (1998)
AJ238250	ndh	NADH dehydrogenase	
AJ238703	porA	Porin	Lichtinger, T. et al. "Biochemical and biophysical characterization of the cell wall porin of <i>Corynebacterium glutamicum</i> : The channel is formed by a low molecular mass polypeptide," <i>Biochemistry</i> , 37(43):15024-15032 (1998)
D17429		Transposable element IS31831	Vertes et al. "Isolation and characterization of IS31831, a transposable element from <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 11(4):739-746 (1994)
D84102	odhA	2-oxoglutarate dehydrogenase	Usuda, Y. et al. "Molecular cloning of the <i>Corynebacterium glutamicum</i> (Brevibacterium lactofermentum AJ12036) odhA gene encoding a novel type of 2-oxoglutarate dehydrogenase," <i>Microbiology</i> , 142:3347-3354 (1996)
E01358	hdh; hk	Homoserine dehydrogenase; homoserine kinase	Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 1 10/12/87
E01359		Upstream of the start codon of homoserine kinase gene	Katsumata, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 2 10/12/87
E01375		Tryptophan operon	
E01376	trpL; trpE	Leader peptide; anthranilate synthase	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87

Table 2 (continued)

	Promoter and operator regions of tryptophan operon	
E01377		Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87
E03937	Biotin-synthase	Hatakeyama, K. et al. "DNA fragment containing gene capable of coding biotin synthetase and its utilization," Patent: JP 1992278088-A 1 10/02/92
E04040	Diamino pelargonic acid aminotransferase	Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92
E04041	Desthiobiotinsynthetase	Kohama, K. et al. "Gene coding diaminopelargonic acid aminotransferase and desthiobiotin synthetase and its utilization," Patent: JP 1992330284-A 1 11/18/92
E04307	Flavum aspartase	Kurusu, Y. et al. "Gene DNA coding aspartase and utilization thereof," Patent: JP 1993030977-A 1 02/09/93
E04376	Isocitric acid lyase	Katsumata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93
E04377	Isocitric acid lyase N-terminal fragment	Katsumata, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993056782-A 3 03/09/93
E04484	Prephenate dehydratase	Sotouchi, N. et al. "Production of L-phenylalanine by fermentation," Patent: JP 1993076352-A 2 03/30/93
E05108	Aspartokinase	Fugono, N. et al. "Gene DNA coding Aspartokinase and its use," Patent: JP 1993184366-A 1 07/27/93
E05112	Dihydro-dipichorinate synthetase	Hatakeyama, K. et al. "Gene DNA coding dihydrodipicolinic acid synthetase and its use," Patent: JP 1993184371-A 1 07/27/93
E05776	Diaminopimelic acid dehydrogenase	Kobayashi, M. et al. "Gene DNA coding Diaminopimelic acid dehydrogenase and its use," Patent: JP 1993284970-A 1 11/02/93
E05779	Threonine synthase	Kohama, K. et al. "Gene DNA coding threonine synthase and its use," Patent: JP 1993284972-A 1 11/02/93
E06110	Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93
E06111	Mutated Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent: JP 1993344881-A 1 12/27/93
E06146	Acetohydroxy acid synthetase	Inui, M. et al. "Gene capable of coding Acetohydroxy acid synthetase and its use," Patent: JP 1993344893-A 1 12/27/93
E06825	Aspartokinase	Sugimoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94
E06826	Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94

Table 2 (continued)

E06827		Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent: JP 1994062866-A 1 03/08/94
E07701	secY		Honno, N. et al. "Gene DNA participating in integration of membrane protein to membrane," Patent: JP 1994169780-A 1 06/21/94
E08177		Aspartokinase	Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94
E08178, E08179, E08180, E08181, E08182		Feedback inhibition-released Aspartokinase	Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent: JP 1994261766-A 1 09/20/94
E08232		Acetohydroxy-acid isomeroreductase	Inui, M. et al. "Gene DNA coding acetohydroxy acid isomeroreductase," Patent: JP 1994277067-A 1 10/04/94
E08234	secE		Asai, Y. et al. "Gene DNA coding for translocation machinery of protein," Patent: JP 1994277073-A 1 10/04/94
E08643		FT aminotransferase and desthiobiotin synthetase promoter region	Hatakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031476-A 1 02/03/95
E08646		Biotin synthetase	Hatakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031476-A 1 02/03/95
E08649		Aspartase	Kohama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent: JP 1995031478-A 1 02/03/95
E08900		Dihydrodipicolinate reductase	Madori, M. et al. "DNA fragment containing gene coding Dihydrodipicolinate acid reductase and utilization thereof," Patent: JP 1995075578-A 1 03/20/95
E08901		Diaminopimelic acid decarboxylase	Madori, M. et al. "DNA fragment containing gene coding Diaminopimelic acid decarboxylase and utilization thereof," Patent: JP 1995075579-A 1 03/20/95
E12594		Serine hydroxymethyltransferase	Hatakeyama, K. et al. "Production of L-tryptophan," Patent: JP 1997028391-A 1 02/04/97
E12760, E12759, E12758		transposase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12764		Arginyl-tRNA synthetase; diaminopimelic acid decarboxylase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12767		Dihydrodipicolinic acid synthetase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12770		aspartokinase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97
E12773		Dihydrodipicolinic acid reductase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent: JP 1997070291-A 03/18/97

Table 2 (continued)

E13655		Glucose-6-phosphate dehydrogenase	Hatakeyama, K. et al. "Glucose-6-phosphate dehydrogenase and DNA capable of coding the same," Patent: JP 1997224661-A 1 09/02/97
L01508	ilvA	Threonine dehydratase	Moeckel, B. et al. "Functional and structural analysis of the threonine dehydratase of <i>Corynebacterium glutamicum</i> ," <i>J. Bacteriol.</i> , 174:8065-8072 (1992)
L07603	EC 4.2.1.15	3-deoxy-D-arabinoheptulosonate-7-phosphate synthase	Chen, C. et al. "The cloning and nucleotide sequence of <i>Corynebacterium glutamicum</i> 3-deoxy-D-arabinoheptulosonate-7-phosphate synthase gene," <i>FEMS Microbiol. Lett.</i> , 107:223-230 (1993)
L09232	ilvB; ilvN; ilvC	Acetohydroxy acid synthase large subunit; Acetohydroxy acid synthase small subunit; Acetohydroxy acid isomeroreductase	Keilhauer, C. et al. "Isolation and synthesis in <i>Corynebacterium glutamicum</i> : molecular analysis of the <i>ilvB-ilvN-ilvC</i> operon," <i>J. Bacteriol.</i> , 175(17):5595-5603 (1993)
L18874	PtsM	Phosphoenolpyruvate sugar phosphotransferase	Fouet, A. et al. "Bacillus subtilis sucrose-specific enzyme II of the phosphotransferase system: expression in <i>Escherichia coli</i> and homology to enzymes II from enteric bacteria," <i>PNAS USA</i> , 84(24):8773-8777 (1987); Lee, J.K. et al. "Nucleotide sequence of the gene encoding the <i>Corynebacterium glutamicum</i> mannose enzyme II and analyses of the deduced protein sequence," <i>FEMS Microbiol. Lett.</i> , 119(1-2):137-145 (1994)
L27123	aceB	Malate synthase	Lee, H-S. et al. "Molecular characterization of <i>aceB</i> , a gene encoding malate synthase in <i>Corynebacterium glutamicum</i> ," <i>J. Microbiol. Biotechnol.</i> , 4(4):256-263 (1994)
L27126		Pyruvate kinase	Jetten, M. S. et al. "Structural and functional analysis of pyruvate kinase from <i>Corynebacterium glutamicum</i> ," <i>Appl. Environ. Microbiol.</i> , 60(7):2501-2507 (1994)
L28760	aceA	Isocitrate lyase	
L35906	dxr	Diphtheria toxin repressor	Oguiza, J.A. et al. "Molecular cloning, DNA sequence analysis, and characterization of the <i>Corynebacterium diphtheriae</i> <i>dxr</i> from <i>Brevibacterium lactofermentum</i> ," <i>J. Bacteriol.</i> , 177(2):465-467 (1995)
M13774		Prephenate dehydratase	Follettie, M.T. et al. "Molecular cloning and nucleotide sequence of the <i>Corynebacterium glutamicum</i> <i>pheA</i> gene," <i>J. Bacteriol.</i> , 167:695-702 (1986)
M16175	5S rRNA		Park, Y-H. et al. "Phylogenetic analysis of the coryneform bacteria by 5S rRNA sequences," <i>J. Bacteriol.</i> , 169:1801-1806 (1987)
M16663	trpE	Anthranelate synthase, 5' end	Sano, K. et al. "Structure and function of the <i>trp</i> operon control regions of <i>Brevibacterium lactofermentum</i> , a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987)
M16664	trpA	Tryptophan synthase, 3' end	Sano, K. et al. "Structure and function of the <i>trp</i> operon control regions of <i>Brevibacterium lactofermentum</i> , a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987)

Table 2 (continued)

M25819		Phosphoenolpyruvate carboxylase	O'Regan, M. et al. "Cloning and nucleotide sequence of the Phosphoenolpyruvate carboxylase-coding gene of <i>Corynebacterium glutamicum</i> ATCC13032," <i>Gene</i> , 77(2):237-251 (1989)
M85106		23S rRNA gene insertion sequence	Roller, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992)
M85107, M85108		23S rRNA gene insertion sequence	Roller, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992)
M89931	aecD; brnQ; yhbW	Beta C-S lyase; branched-chain amino acid uptake carrier; hypothetical protein yhbW	Rosol, I. et al. "The <i>Corynebacterium glutamicum</i> aecD gene encodes a C-S lyase with alpha, beta-elimination activity that degrades aminoethylcysteine," <i>J. Bacteriol.</i> , 174(9):2968-2977 (1992); Tauch, A. et al. "Isoleucine uptake in <i>Corynebacterium glutamicum</i> ATCC 13032 is directed by the brnQ gene product," <i>Arch. Microbiol.</i> , 169(4):303-312 (1998)
S59299	trp	Leader gene (promoter)	Herry, D.M. et al. "Cloning of the trp gene cluster from a tryptophan-hyperproducing strain of <i>Corynebacterium glutamicum</i> : identification of a mutation in the trp leader sequence," <i>Appl. Environ. Microbiol.</i> , 59(3):791-799 (1993)
U11545	trpD	Anthranilate phosphoribosyltransferase	O'Gara, J.P. and Dunican, L.K. (1994) Complete nucleotide sequence of the <i>Corynebacterium glutamicum</i> ATCC 21850 trpD gene." Thesis, Microbiology Department, University College Galway, Ireland.
U13922	cgIIIM; cgIIIR; cglIIR	Putative type II 5-cytosine methyltransferase; putative type II restriction endonuclease; putative type I or type III restriction endonuclease	Schafer, A. et al. "Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from <i>Corynebacterium glutamicum</i> ATCC 13032 and analysis of its role in intergeneric conjugation with <i>Escherichia coli</i> ," <i>J. Bacteriol.</i> , 176(23):7309-7319 (1994); Schafer, A. et al. "The <i>Corynebacterium glutamicum</i> cglIIM gene encoding a 5-cytosine in an McrBC-deficient <i>Escherichia coli</i> strain," <i>Gene</i> , 203(2):95-101 (1997)
U14965	recA		
U31224	ppx		Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31225	proC	L-proline: NADP+ 5-oxido-reductase	Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31230	obg; proB; unkdh	?gamma glutamyl kinase; similar to D-isomer specific 2-hydroxyacid dehydrogenases	Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)

Table 2 (continued)

	bioB	Biotin synthase	
U31281			Serebriiskii, I.G., "Two new members of the bio B superfamily: Cloning, sequencing and expression of bio B genes of <i>Methylobacillus flagellatum</i> and <i>Corynebacterium glutamicum</i> ," <i>Gene</i> , 175:15-22 (1996)
U35023	thtR; accBC	Thiosulfate sulfurtransferase; acyl CoA carboxylase	Jager, W. et al. "A <i>Corynebacterium glutamicum</i> gene encoding a two-domain protein similar to biotin carboxylases and biotin-carboxyl-carrier proteins," <i>Arch. Microbiol.</i> , 166(2):76-82 (1996)
U43535	cmr	Multidrug resistance protein	Jager, W. et al. "A <i>Corynebacterium glutamicum</i> gene conferring multidrug resistance in the heterologous host <i>Escherichia coli</i> ," <i>J. Bacteriol.</i> , 179(7):2449-2451 (1997)
U43536	clpB	Heat shock ATP-binding protein	
U53587	aphA-3	3'5'-aminoglycoside phosphotransferase	
U89648		<i>Corynebacterium glutamicum</i> unidentified sequence involved in histidine biosynthesis; partial sequence	
X04960	trpA; trpB; trpC; trpD; trpE; trpG; trpL	Tryptophan operon	Matsui, K. et al. "Complete nucleotide and deduced amino acid sequences of the <i>Brevibacterium lactofermentum</i> tryptophan operon," <i>Nucleic Acids Res.</i> , 14(24):10113-10114 (1986)
X07563	lys A	DAP decarboxylase (meso-diaminopimelate decarboxylase, EC 4.1.1.20)	Yeh, P. et al. "Nucleic sequence of the lysA gene of <i>Corynebacterium glutamicum</i> and possible mechanisms for modulation of its expression," <i>Mol. Gen. Genet.</i> , 212(1):112-119 (1988)
X14234	EC 4.1.1.31	Phosphoenolpyruvate carboxylase	Eikmanns, B.J. et al. "The Phosphoenolpyruvate carboxylase gene of <i>Corynebacterium glutamicum</i> : Molecular cloning, nucleotide sequence, and expression," <i>Mol. Gen. Genet.</i> , 218(2):330-339 (1989); Lepiniec, L. et al. "Sorghum Phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution," <i>Plant. Mol. Biol.</i> , 21 (3):487-502 (1993)
X17313	fda	Fructose-bisphosphate aldolase	Von der Osten, C.H. et al. "Molecular cloning, nucleotide sequence and fine-structural analysis of the <i>Corynebacterium glutamicum</i> fda gene: structural comparison of <i>C. glutamicum</i> fructose-1, 6-bisphosphate aldolase to class I and class II aldolases," <i>Mol. Microbiol.</i> ,
X53993	dapA	L-2, 3-dihydrodipicolinate synthetase (EC 4.2.1.52)	Bonnassie, S. et al. "Nucleic sequence of the dapA gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> , 18(21):6421 (1990)
X54223		AttB-related site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of <i>Corynebacterium diphtheriae</i> , <i>Corynebacterium ulcerans</i> , <i>Corynebacterium glutamicum</i> , and the attP site of <i>lambdacorynebophage</i> ," <i>FEBS. Microbiol. Lett.</i> , 66:299-302 (1990)
X54740	argS; lysA	Arginyl-tRNA synthetase; Diaminopimelate decarboxylase	Marcel, T. et al. "Nucleotide sequence and organization of the upstream region of the <i>Corynebacterium glutamicum</i> lysA gene," <i>Mol. Microbiol.</i> , 4(11):1819-1830 (1990)

Table 2 (continued)

X55994	trpL; trpE	Putative leader peptide; anthranilate synthase component 1	Heery, D.M. et al. "Nucleotide sequence of the <i>Corynebacterium glutamicum</i> trpE gene," <i>Nucleic Acids Res.</i> , 18(23):7138 (1990)
X56037	thrC	Threonine synthase	Han, K.S. et al. "The molecular structure of the <i>Corynebacterium glutamicum</i> threonine synthase gene," <i>Mol. Microbiol.</i> , 4(10):1693-1702 (1990)
X56075	attB-related site	Attachment site	Ciancioffo, N. et al. "DNA sequence homology between att B-related sites of <i>Corynebacterium diphtheriae</i> , <i>Corynebacterium ulcerans</i> , <i>Corynebacterium glutamicum</i> , and the attP site of <i>lambdacorynephage</i> ," <i>FEMS. Microbiol. Lett.</i> , 66:299-302 (1990)
X57226	lysC-alpha; lysC-beta; asd	Aspartokinase-alpha subunit; Aspartokinase-beta subunit; aspartate beta semialdehyde dehydrogenase	Kalinowski, J. et al. "Genetic and biochemical analysis of the Aspartokinase from <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 5(5):1197-1204 (1991); Kalinowski, J. et al. "Aspartokinase genes lysC alpha and lysC beta overlap and are adjacent to the aspartate beta-semialdehyde dehydrogenase gene asd in <i>Corynebacterium glutamicum</i> ," <i>Mol. Gen. Genet.</i> , 224(3):317-324 (1990)
X59403	gap;pgk; tpi	Glyceraldehyde-3-phosphate; phosphoglycerate kinase; triosephosphate isomerase	Eikmanns, B.J. "Identification, sequence analysis, and expression of a <i>Corynebacterium glutamicum</i> gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase," <i>J. Bacteriol.</i> , 174(19):6076-6086 (1992)
X59404	gdh	Glutamate dehydrogenase	Bormann, E.R. et al. "Molecular analysis of the <i>Corynebacterium glutamicum</i> gdh gene encoding glutamate dehydrogenase," <i>Mol. Microbiol.</i> , 6(3):317-326 (1992)
X60312	lysI	L-lysine permease	Seep-Feldhaus, A.H. et al. "Molecular analysis of the <i>Corynebacterium glutamicum</i> lysI gene involved in lysine uptake," <i>Mol. Microbiol.</i> , 5(12):2995-3005 (1991)
X66078	copI	PsI protein	Joliff, G. et al. "Cloning and nucleotide sequence of the cspI gene encoding PS1, one of the two major secreted proteins of <i>Corynebacterium glutamicum</i> : The deduced N-terminal region of PS1 is similar to the <i>Mycobacterium</i> antigen 85 complex," <i>Mol. Microbiol.</i> , 6(16):2349-2362 (1992)
X66112	glt	Citrate synthase	Eikmanns, B.J. et al. "Cloning sequence, expression and transcriptional analysis of the <i>Corynebacterium glutamicum</i> gltA gene encoding citrate synthase," <i>Microbiol.</i> , 140:1817-1828 (1994)
X67737	dapB	Dihydrodipicolinate reductase	
X69103	csp2	Surface layer protein PS2	Peyret, J.L. et al. "Characterization of the cspB gene encoding PS2, an ordered surface-layer protein in <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 9(1):97-109 (1993)
X69104		IS3 related insertion element	Bonamy, C. et al. "Identification of IS1206, a <i>Corynebacterium glutamicum</i> IS3-related insertion sequence and phylogenetic analysis," <i>Mol. Microbiol.</i> , 14(3):571-581 (1994)

Table 2 (continued)

X70959	leuA	Isopropylmalate synthase	Patek, M. et al. "Leucine synthesis in Corynebacterium glutamicum: enzyme activities, structure of leuA, and effect of leuA inactivation on lysine synthesis," <i>Appl. Environ. Microbiol.</i> , 60(1):133-140 (1994)
X71489	icd	Isocitrate dehydrogenase (NADP+)	Eikmanns, B.J. et al. "Cloning sequence analysis, expression, and inactivation of the Corynebacterium glutamicum icd gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme," <i>J. Bacteriol.</i> , 177(3):774-782 (1995)
X72855	GDHA	Glutamate dehydrogenase (NADP+)	Heery, D.M. et al. "A sequence from a tryptophan-hyperproducing strain of Corynebacterium glutamicum encoding resistance to 5-methyltryptophan," <i>Biochem. Biophys. Res. Commun.</i> , 201(3):1255-1262 (1994)
X75083, X70584	mtrA	5-methyltryptophan resistance	Fitzpatrick, R. et al. "Construction and characterization of recA mutant strains of Corynebacterium glutamicum and Brevibacterium lactofermentum," <i>Appl. Microbiol. Biotechnol.</i> , 42(4):575-580 (1994)
X75085	recA		Reinscheid, D.J. et al. "Characterization of the isocitrate lyase gene from Corynebacterium glutamicum and biochemical analysis of the enzyme," <i>J. Bacteriol.</i> , 176(12):3474-3483 (1994)
X75504	aceA; thiX	Partial isocitrate lyase; ?	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993)
X76875		ATPase beta-subunit	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993)
X77034	tuf	Elongation factor Tu	Billman-Jacobe, H. "Nucleotide sequence of a recA gene from Corynebacterium glutamicum," <i>DNA Seq.</i> , 4(6):403-404 (1994)
X77384	recA		Reinscheid, D.J. et al. "Malate synthase from Corynebacterium glutamicum pta-ack operon encoding phosphotransacetylase: sequence analysis," <i>Microbiology</i> , 140:3099-3108 (1994)
X78491	aceB	Malate synthase	Rainey, F.A. et al. "Phylogenetic analysis of the genera Rhodococcus and Norcardia and evidence for the evolutionary origin of the genus Norcardia from within the radiation of Rhodococcus species," <i>Microbiol.</i> , 141:523-528 (1995)
X80629	16S rDNA	16S ribosomal RNA	Kronmeyer, W. et al. "Structure of the gluABCD cluster encoding the glutamate uptake system of Corynebacterium glutamicum," <i>J. Bacteriol.</i> , 177(5):1152-1158 (1995)
X81191	gluA; gluB; gluC; gluD	Glutamate uptake system	Wehrmann, A. et al. "Analysis of different DNA fragments of Corynebacterium glutamicum complementing dapE of Escherichia coli," <i>Microbiology</i> , 40:3349-56 (1994)
X81379	dapE	Succinyl-diaminopimelate desuccinylase	



Table 2 (continued)

	16S rDNA	16S ribosomal RNA	
X82061			Ruimy, R. et al. "Phylogeny of the genus <i>Corynebacterium</i> deduced from analyses of small-subunit ribosomal DNA sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):740-746 (1995)
X82928	asd; lysC	Aspartate-semialdehyde dehydrogenase; ?	Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995)
X82929	proA	Gamma-glutamyl phosphate reductase	Serebrijski, I. et al. "Multicopy suppression by asd gene and osmotic stress-dependent complementation by heterologous proA in proA mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995)
X84257	16S rDNA	16S ribosomal RNA	Pascual, C. et al. "Phylogenetic analysis of the genus <i>Corynebacterium</i> based on 16S rRNA gene sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):724-728 (1995)
X85965	aroP; dapE	Aromatic amino acid permease; ?	Wehrmann et al. "Functional analysis of sequences adjacent to dapE of <i>C. glutamicum</i> proline reveals the presence of aroP, which encodes the aromatic amino acid transporter," <i>J. Bacteriol.</i> , 177(20):5991-5993 (1995)
X86157	argB; argC; argD; argF; argJ	Acetylglutamate kinase; N-acetyl-gamma-glutamyl-phosphate reductase; acetylornithine aminotransferase; ornithine carbamoyltransferase; glutamate N-acetyltransferase	Sakanyan, V. et al. "Genes and enzymes of the acetyl cycle of arginine biosynthesis in <i>Corynebacterium glutamicum</i> : enzyme evolution in the early steps of the arginine pathway," <i>Microbiology</i> , 142:99-108 (1996)
X89084	pta; ackA	Phosphate acetyltransferase; acetate kinase	Reinscheid, D.J. et al. "Cloning, sequence analysis, expression and inactivation of the <i>Corynebacterium glutamicum</i> pta-ack operon encoding phosphotransacetylase and acetate kinase," <i>Microbiology</i> , 145:503-513 (1999)
X89850	attB	Attachment site	Le Marrec, C. et al. "Genetic characterization of site-specific integration functions of phi AAU2 infecting 'Arthrobacter aureus C70,'" <i>J. Bacteriol.</i> , 178(7):1996-2004 (1996)
X90356		Promoter fragment F1	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90357		Promoter fragment F2	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90358		Promoter fragment F10	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90359		Promoter fragment F13	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)

Table 2 (continued)

X90360	Promoter fragment F22	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90361	Promoter fragment F34	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90362	Promoter fragment F37	Patek, M. et al. "Promoters from <i>C. glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90363	Promoter fragment F45	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90364	Promoter fragment F64	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90365	Promoter fragment F75	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90366	Promoter fragment PF101	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90367	Promoter fragment PF104	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X90368	Promoter fragment PF109	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> , 142:1297-1309 (1996)
X93513	Ammonium transport system	Siewe, R.M. et al. "Functional and genetic characterization of the (methyl) ammonium uptake carrier of <i>Corynebacterium glutamicum</i> ," <i>J. Biol. Chem.</i> , 271(10):5398-5403 (1996)
X93514	Glycine betaine transport system	Peter, H. et al. "Isolation, characterization, and expression of the <i>Corynebacterium glutamicum</i> betP gene, encoding the transport system for the compatible solute glycine betaine," <i>J. Bacteriol.</i> , 178(17):5229-5234 (1996)
X95649	orf4	Patek, M. et al. "Identification and transcriptional analysis of the dapB-ORF2-dapA-ORF4 operon of <i>Corynebacterium glutamicum</i> , encoding two enzymes involved in L-lysine synthesis," <i>Biotechnol. Lett.</i> , 19:1113-1117 (1997)
X96471	Lysine exporter protein; Lysine export regulator protein	Vrijic, M. et al. "A new type of transporter with a new type of cellular function: L-lysine export from <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 22(5):815-826 (1996)

Table 2 (continued)

X96580	panB; panC; xylB	3-methyl-2-oxobutanate hydroxymethyltransferase; pantoate-beta-alanine ligase; xylulokinase	Sahn, H. et al. "D-pantothenate synthesis in <i>Corynebacterium glutamicum</i> and use of panBC and genes encoding L-valine synthesis for D-pantothenate overproduction," <i>Appl. Environ. Microbiol.</i> , 65(5):1973-1979 (1999)
X96962		Insertion sequence IS1207 and transposase	
X99289		Elongation factor P	Ramos, A. et al. "Cloning, sequencing and expression of the gene encoding elongation factor P in the amino-acid producer <i>Brevibacterium lactofermentum</i> ( <i>Corynebacterium glutamicum</i> ATCC 13869)," <i>Gene</i> , 198:217-222 (1997)
Y00140	thrB	Homoserine kinase	Mateos, L.M. et al. "Nucleotide sequence of the homoserine kinase (thrB) gene of the <i>Brevibacterium lactofermentum</i> ," <i>Nucleic Acids Res.</i> , 15(9):3922 (1987)
Y00151	ddh	Meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16)	Ishino, S. et al. "Nucleotide sequence of the meso-diaminopimelate D-dehydrogenase gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> , 15(9):3917 (1987)
Y00476	thrA	Homoserine dehydrogenase	Mateos, L.M. et al. "Nucleotide sequence of the homoserine dehydrogenase (thrA) gene of the <i>Brevibacterium lactofermentum</i> ," <i>Nucleic Acids Res.</i> , 15(24):10598 (1987)
Y00546	hom; thrB	Homoserine dehydrogenase; homoserine kinase	Peoples, O.P. et al. "Nucleotide sequence and fine structural analysis of the <i>Corynebacterium glutamicum</i> hom-thrB operon," <i>Mol. Microbiol.</i> , 2(1):63-72 (1988)
Y08964	murC; ftsQ/divD; ftsZ	UPD-N-acetylmuramate-alanine ligase; division initiation protein or cell division protein; cell division protein	Honrubia, M.P. et al. "Identification, characterization, and chromosomal organization of the ftsZ gene from <i>Brevibacterium lactofermentum</i> ," <i>Mol. Gen. Genet.</i> , 259(1):97-104 (1998)
Y09163	putP	High affinity proline transport system	Peter, H. et al. "Isolation of the putP gene of <i>Corynebacterium glutamicum</i> proline and characterization of a low-affinity uptake system for compatible solutes," <i>Arch. Microbiol.</i> , 168(2):143-151 (1997)
Y09548	pyc	Pyruvate carboxylase	Peters-Wendisch, P.G. et al. "Pyruvate carboxylase from <i>Corynebacterium glutamicum</i> : characterization, expression and inactivation of the pyc gene," <i>Microbiology</i> , 144:915-927 (1998)
Y09578	leuB	3-isopropylmalate dehydrogenase	Patek, M. et al. "Analysis of the leuB gene from <i>Corynebacterium glutamicum</i> ," <i>Appl. Microbiol. Biotechnol.</i> , 50(1):42-47 (1998)
Y12472		Attachment site bacteriophage Phi-16	Moreau, S. et al. "Site-specific integration of corynephage Phi-16: The construction of an integration vector," <i>Microbiol.</i> , 145:539-548 (1999)
Y12537	proP	Proline/ectoine uptake system protein	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes: Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EctP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998)

Table 2 (continued)

Y13221	glnA	Glutamine synthetase I	Jakoby, M. et al. "Isolation of <i>Corynebacterium glutamicum</i> glnA gene encoding glutamine synthetase I," <i>FEMS Microbiol. Lett.</i> , 154(1):81-88 (1997)
Y16642	lpd	Dihydrolipoamide dehydrogenase	
Y18059		Attachment site <i>Corynebacterium</i> 304L	
Z21501	argS; lysA	Arginyl-tRNA synthetase; diaminopimelate decarboxylase (partial)	Moreau, S. et al. "Analysis of the integration functions of $\phi$ phi304L: An integrase module among corynebacteriophages," <i>Virology</i> , 255(1):150-159 (1999)
Z21502	dapA; dapB	Dihydrodipicolinate synthase; dihydrodipicolinate reductase	Oguiza, J.A. et al. "A gene encoding arginyl-tRNA synthetase is located in the upstream region of the lysA gene in <i>Brevibacterium lactofermentum</i> : Regulation of argS-lysA cluster expression by arginine," <i>J. Bacteriol.</i> , 175(22):7356-7362 (1993)
Z29563	thrC	Threonine synthase	Pisabarro, A. et al. "A cluster of three genes (dapA, orf2, and dapB) of <i>Brevibacterium lactofermentum</i> encodes dihydrodipicolinate reductase, and a third polypeptide of unknown function," <i>J. Bacteriol.</i> , 175(9):2743-2749 (1993)
Z46753	16S rDNA	Gene for 16S ribosomal RNA	Malumbres, M. et al. "Analysis and expression of the thrC gene of the encoded threonine synthase," <i>Appl. Environ. Microbiol.</i> , 60(7):2209-2219 (1994)
Z49822	sigA	SigA sigma factor	Oguiza, J.A. et al. "Multiple sigma factor genes in <i>Brevibacterium lactofermentum</i> : Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996)
Z49823	galE; dxr	Catalytic activity UDP-galactose 4-epimerase; diphtheria toxin regulatory protein	Oguiza, J.A. et al. "The galE gene encoding the UDP-galactose 4-epimerase of <i>Brevibacterium lactofermentum</i> is coupled transcriptionally to the dmdR gene," <i>Gene</i> , 177:103-107 (1996)
Z49824	orf1; sigB	?; SigB sigma factor	Oguiza, J.A. et al. "Multiple sigma factor genes in <i>Brevibacterium lactofermentum</i> : Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996)
Z66534		Transposase	Correia, A. et al. "Cloning and characterization of an IS-like element present in the genome of <i>Brevibacterium lactofermentum</i> ATCC 13869," <i>Gene</i> , 170(1):91-94 (1996)

\* A sequence for this gene was published in the indicated reference. However, the sequence obtained by the inventors of the present application is significantly longer than the published version. It is believed that the published version relied on an incorrect start codon, and thus represents only a fragment of the actual coding region.

TABLE 3: *Corynebacterium* and *Brevibacterium* Strains Which May be Used in the Practice of the Invention

Genus	species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
<i>Brevibacterium</i>	<i>ammoniagenes</i>	21054							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19350							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19351							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19352							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19353							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19354							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19355							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	19356							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	21055							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	21077							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	21553							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	21580							
<i>Brevibacterium</i>	<i>ammoniagenes</i>	39101							
<i>Brevibacterium</i>	<i>butanicum</i>	21196							
<i>Brevibacterium</i>	<i>divaricatum</i>	21792	P928						
<i>Brevibacterium</i>	<i>flavum</i>	21474							
<i>Brevibacterium</i>	<i>flavum</i>	21129							
<i>Brevibacterium</i>	<i>flavum</i>	21518							
<i>Brevibacterium</i>	<i>flavum</i>			B11474					
<i>Brevibacterium</i>	<i>flavum</i>			B11472					
<i>Brevibacterium</i>	<i>flavum</i>	21127							
<i>Brevibacterium</i>	<i>flavum</i>	21128							
<i>Brevibacterium</i>	<i>flavum</i>	21427							
<i>Brevibacterium</i>	<i>flavum</i>	21475							
<i>Brevibacterium</i>	<i>flavum</i>	21517							
<i>Brevibacterium</i>	<i>flavum</i>	21528							
<i>Brevibacterium</i>	<i>flavum</i>	21529							
<i>Brevibacterium</i>	<i>flavum</i>			B11477					
<i>Brevibacterium</i>	<i>flavum</i>			B11478					
<i>Brevibacterium</i>	<i>flavum</i>	21127							
<i>Brevibacterium</i>	<i>flavum</i>			B11474					
<i>Brevibacterium</i>	<i>healii</i>	15527							
<i>Brevibacterium</i>	<i>ketoglutamicum</i>	21004							
<i>Brevibacterium</i>	<i>ketoglutamicum</i>	21089							
<i>Brevibacterium</i>	<i>ketosoreductum</i>	21914							
<i>Brevibacterium</i>	<i>lactofermentum</i>				70				
<i>Brevibacterium</i>	<i>lactofermentum</i>				74				
<i>Brevibacterium</i>	<i>lactofermentum</i>				77				
<i>Brevibacterium</i>	<i>lactofermentum</i>	21798							
<i>Brevibacterium</i>	<i>lactofermentum</i>	21799							
<i>Brevibacterium</i>	<i>lactofermentum</i>	21800							
<i>Brevibacterium</i>	<i>lactofermentum</i>	21801							
<i>Brevibacterium</i>	<i>lactofermentum</i>			B11470					
<i>Brevibacterium</i>	<i>lactofermentum</i>			B11471					

Genus	species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Brevibacterium	lactofermentum	21086							
Brevibacterium	lactofermentum	21420							
Brevibacterium	lactofermentum	21086							
Brevibacterium	lactofermentum	31269							
Brevibacterium	linens	9174							
Brevibacterium	linens	19391							
Brevibacterium	linens	8377							
Brevibacterium	paraffinolyticum					11160			
Brevibacterium	spec.						717.73		
Brevibacterium	spec.						717.73		
Brevibacterium	spec.	14604							
Brevibacterium	spec.	21860							
Brevibacterium	spec.	21864							
Brevibacterium	spec.	21865							
Brevibacterium	spec.	21866							
Brevibacterium	spec.	19240							
Corynebacterium	acetoacidophilum	21476							
Corynebacterium	acetoacidophilum	13870							
Corynebacterium	acetoglutamicum			B11473					
Corynebacterium	acetoglutamicum			B11475					
Corynebacterium	acetoglutamicum	15806							
Corynebacterium	acetoglutamicum	21491							
Corynebacterium	acetoglutamicum	31270							
Corynebacterium	acetophilum			B3671					
Corynebacterium	ammoniagenes	6872						2399	
Corynebacterium	ammoniagenes	15511							
Corynebacterium	fujikense	21496							
Corynebacterium	glutamicum	14067							
Corynebacterium	glutamicum	39137							
Corynebacterium	glutamicum	21254							
Corynebacterium	glutamicum	21255							
Corynebacterium	glutamicum	31830							
Corynebacterium	glutamicum	13032							
Corynebacterium	glutamicum	14305							
Corynebacterium	glutamicum	15455							
Corynebacterium	glutamicum	13058							
Corynebacterium	glutamicum	13059							
Corynebacterium	glutamicum	13060							
Corynebacterium	glutamicum	21492							
Corynebacterium	glutamicum	21513							
Corynebacterium	glutamicum	21526							
Corynebacterium	glutamicum	21543							
Corynebacterium	glutamicum	13287							
Corynebacterium	glutamicum	21851							
Corynebacterium	glutamicum	21253							
Corynebacterium	glutamicum	21514							
Corynebacterium	glutamicum	21516							
Corynebacterium	glutamicum	21299							

Genus	Species	ATCC	FERM	NRRL	CECT	NCIMB	GBS	NCTC	DSMZ
Corynebacterium	glutamicum	21300							
Corynebacterium	glutamicum	39684							
Corynebacterium	glutamicum	21488							
Corynebacterium	glutamicum	21649							
Corynebacterium	glutamicum	21650							
Corynebacterium	glutamicum	19223							
Corynebacterium	glutamicum	13869							
Corynebacterium	glutamicum	21157							
Corynebacterium	glutamicum	21158							
Corynebacterium	glutamicum	21159							
Corynebacterium	glutamicum	21355							
Corynebacterium	glutamicum	31808							
Corynebacterium	glutamicum	21674							
Corynebacterium	glutamicum	21562							
Corynebacterium	glutamicum	21563							
Corynebacterium	glutamicum	21564							
Corynebacterium	glutamicum	21565							
Corynebacterium	glutamicum	21566							
Corynebacterium	glutamicum	21567							
Corynebacterium	glutamicum	21568							
Corynebacterium	glutamicum	21569							
Corynebacterium	glutamicum	21570							
Corynebacterium	glutamicum	21571							
Corynebacterium	glutamicum	21572							
Corynebacterium	glutamicum	21573							
Corynebacterium	glutamicum	21579							
Corynebacterium	glutamicum	19049							
Corynebacterium	glutamicum	19050							
Corynebacterium	glutamicum	19051							
Corynebacterium	glutamicum	19052							
Corynebacterium	glutamicum	19053							
Corynebacterium	glutamicum	19054							
Corynebacterium	glutamicum	19055							
Corynebacterium	glutamicum	19056							
Corynebacterium	glutamicum	19057							
Corynebacterium	glutamicum	19058							
Corynebacterium	glutamicum	19059							
Corynebacterium	glutamicum	19060							
Corynebacterium	glutamicum	19185							
Corynebacterium	glutamicum	13286							
Corynebacterium	glutamicum	21515							
Corynebacterium	glutamicum	21527							
Corynebacterium	glutamicum	21544							
Corynebacterium	glutamicum	21492							
Corynebacterium	glutamicum			B8183					
Corynebacterium	glutamicum			B8182					
Corynebacterium	glutamicum			B12416					
Corynebacterium	glutamicum			B12417					

Genus	species	ATCC	FERM	NRRL	CECT	NCIMB	CBS	NCTC	DSMZ
Corynebacterium	glutamicum			B12418					
Corynebacterium	glutamicum			B11476					
Corynebacterium	glutamicum	21608							
Corynebacterium	lilium		P973						
Corynebacterium	nitrilophilus	21419				11594			
Corynebacterium	spec.		P4445						
Corynebacterium	spec.		P4446						
Corynebacterium	spec.	31088							
Corynebacterium	spec.	31089							
Corynebacterium	spec.	31090							
Corynebacterium	spec.	31090							
Corynebacterium	spec.	31090							
Corynebacterium	spec.	15954							20145
Corynebacterium	spec.	21857							
Corynebacterium	spec.	21862							
Corynebacterium	spec.	21863							

ATCC: American Type Culture Collection, Rockville, MD, USA

FERM: Fermentation Research Institute, Chiba, Japan

NRRL: ARS Culture Collection, Northern Regional Research Laboratory, Peoria, IL, USA

CECT: Coleccion Espanola de Cultivos Tipo, Valencia, Spain

NCIMB: National Collection of Industrial and Marine Bacteria Ltd., Aberdeen, UK

CBS: Centraalbureau voor Schimmelcultures, Baarn, NL

NCTC: National Collection of Type Cultures, London, UK

DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany

For reference see Sugawara, H. et al. (1993) World directory of collections of cultures of microorganisms: Bacteria, fungi and yeasts (4<sup>th</sup> edn), World federation for culture collections world data center on microorganisms, Saimata, Japan.



Table 4: Alignment Results

ID #	length (NT)	length (NT)	Genbank Hit	Accession	Name of Genbank Hit	Source of Genbank Hit	% homology (GAP)	Date of Deposit
rx00026	1509	158405	GB_RO:MMHC310M6	AF109906	Mus musculus MHC class III region RD gene, partial cds; Bf, C2, G9A, NG22, G9, HSP70, HSP70, HSC70i, and snRNP genes, complete cds; G7A gene, partial cds; and unknown genes.	Mus musculus	38,003	10-DEC-1998
		119007	GB_HTG2:AC007029	AC007029	Homo sapiens clone DJ0855F16, *** SEQUENCING IN PROGRESS	Homo sapiens	37,943	7-Apr-99
		119007	GB_HTG2:AC007029	AC007029	***, 1 unordered pieces.	Homo sapiens	37,943	7-Apr-99
rx00072					***, 1 unordered pieces.			
rx00111	1116	2748	GB_BA1:SAUSIGA	M94370	Stigmatella aurantiaca sigma factor (sigA) gene, complete cds.	Stigmatella aurantiaca	40,435	16-Aug-94
		28500	GB_BA1:SC5B8	AL022374	Streptomyces coelicolor cosmid 588.	Streptomyces coelicolor	40,090	22-Apr-98
		9086	GB_BA2:AE001767	AE001767	Thermotoga maritima section 79 of 136 of the complete genome.	Thermotoga maritima	35,091	2-Jun-99
rx00112	1314	418	GB_EST35:AU075536	AU075536	AU075536 Rice shoot Oryza sativa cDNA clone S0028_2Z, mRNA sequence.	Oryza sativa	39,423	7-Jul-99
		647	GB_GSS9:AU157585	AQ157585	nbxb0009B16r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0009B16r, genomic survey sequence.	Oryza sativa	40,867	12-Sep-98
		542	GB_GSS14:AQ510314	AQ510314	nbxb0095O05f CUGI Rice BAC Library Oryza sativa genomic clone nbxb0095O05f, genomic survey sequence.	Oryza sativa	39,372	04-MAY-1999
rx00133	936	38404	GB_BA1:SC2G5	AL035478	Streptomyces coelicolor cosmid 2G5.	Streptomyces coelicolor	41,170	11-Jun-99
		515	GB_EST7:W64291	W64291	md98h12.r1 Soares mouse embryo NbME13.5 (4.5 Mus musculus cDNA clone IMAGE:386087 5' similar to gb:L26528 Mus musculus Rab11b mRNA, complete cds (MOUSE)), mRNA sequence.	Mus musculus	35,306	10-Jun-96
		39594	GB_PR3:AC005624	AC005624	Homo sapiens chromosome 19, cosmid R30017, complete sequence.	Homo sapiens	39,054	6-Sep-98
rx00137	1212	4115	GB_BA2:AF124600	AF124600	Corynebacterium glutamicum chorismate synthase (aroC), shikimate kinase (aroK), and 3-dehydroquinase synthase (aroB) genes, complete cds; and putative cytoplasmic peptidase (pepQ) gene, partial cds.	Corynebacterium glutamicum	99,867	04-MAY-1999
		33818	GB_BA1:MTCY159	Z83863	Mycobacterium tuberculosis H37Rv complete genome; segment 111/162.	Mycobacterium tuberculosis	40,959	17-Jun-98
		3437	GB_BA1:MT3DEHQ	X59509	M.tuberculosis, genes for 3-dehydroquinase synthase and 3-dehydroquinase.	Mycobacterium tuberculosis	52,583	30-Jun-93
rx00139	834	738	GB_BA1:BLELONP	X99289	B.lactofermentum gene encoding elongation factor P.	Corynebacterium glutamicum	100,000	1-Nov-97
		38666	GB_PL1:SPAC24C9	Z98601	S.pombe chromosome I cosmid c24C9.	Schizosaccharomyces pombe	35,230	24-Feb-99
		1110000	GB_HTG1:CEY102A5	Z99711	Caenorhabditis elegans chromosome V clone Y102A5, *** SEQUENCING IN PROGRESS *** in unordered pieces.	Caenorhabditis elegans	37,775	Z89711

Table 4 (continued)

rx00152	1419	GB_BA1:MTCY277	38300	Z79701	Mycobacterium tuberculosis H37Rv complete genome; segment 65/162.	Mycobacterium tuberculosis	58,500	17-Jun-98
		GB_BA1:MSGY456	37316	AD000001	Mycobacterium tuberculosis sequence from clone y456.	Mycobacterium tuberculosis	38,913	03-DEC-1996
		GB_BA2:AF002133	15437	AF002133	Mycobacterium avium strain GIR10 transcriptional regulator (mav81) gene, partial cds, aconitase (acn), invasin 1 (inv1), invasin 2 (inv2), transcriptional regulator (moxR), ketoacyl-reductase (fabG), enoyl-reductase (inhA) and ferrochelatase (mav272) genes, complete cds.	Mycobacterium avium	64,009	26-MAR-1998
rx00226	948	GB_PR3:AC005756	43299	AC005756	Homo sapiens chromosome 19, fosmid 39347, complete sequence.	Homo sapiens	36,209	02-OCT-1998
		GB_GSS5:AQ818463	413	AQ818463	HS_525Q_A2_B08_SP6E RPC1-11 Human Male BAC Library Homo sapiens genomic clone Plate=826 Col=16 Row=C, genomic survey sequence.	Homo sapiens	37,288	26-Aug-99
		GB_GSS5:AQ782337	832	AQ782337	HS_3184_B1_H12_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3184 Col=23 Row=P, genomic survey sequence.	Homo sapiens	35,917	2-Aug-99
rx00249	980	GB_BA2:AF035608	3614	AF035608	Pseudomonas aeruginosa ATP sulfurylase small subunit (cysD) and ATP sulfurylase GTP-binding subunit/APS kinase (cysN) genes, complete cds.	Pseudomonas aeruginosa	50,205	1-Jun-98
		GB_BA1:AB017641	17101	AB017641	Micromonospora griseorubida gene for polyketide synthase, complete cds.	Micromonospora griseorubida	40,266	2-Apr-99
		GB_BA2:AF002133	15437	AF002133	Mycobacterium avium strain GIR10 transcriptional regulator (mav81) gene, partial cds, aconitase (acn), invasin 1 (inv1), invasin 2 (inv2), transcriptional regulator (moxR), ketoacyl-reductase (fabG), enoyl-reductase (inhA) and ferrochelatase (mav272) genes, complete cds.	Mycobacterium avium	38,429	26-MAR-1998
rx00299	1101	GB_BA2:CORCSLYS	2821	M89931	Corynebacterium glutamicum beta C-S lyase (aecD) and branched-chain amino acid uptake carrier (brnQ) genes, complete cds, and hypothetical protein YhbW (yhbW) gene, partial cds.	Corynebacterium glutamicum	100,000	4-Jun-98
		GB_BA1:CGECTP	2719	AJ001436	Corynebacterium glutamicum ecfP gene.	Corynebacterium glutamicum	41,143	20-Nov-98
		GB_BA2:AF181035	5922	AF181035	Rhodobacter sphaeroides glycogen utilization operon, complete sequence.	Rhodobacter sphaeroides	36,701	7-Sep-99
rx00332	825	GB_BA1:CGTHRC	3120	X56037	Corynebacterium glutamicum thrC gene for threonine synthase (EC 4.2.99.2).	Corynebacterium glutamicum	37,730	17-Jun-97
		GB_PAT:109078	3146	109078	Sequence 4 from Patent WO 8809819.	Unknown.	38,700	02-DEC-1994
		GB_PR3:HSJ333B15	73666	AL109954	Human DNA sequence from clone 333B15 on chromosome 20, complete sequence.	Homo sapiens	37,203	23-Nov-99
rx00470	1392	GB_PL2:DCPCNAM	865	X62977	D. carota mRNA for proliferating cell nuclear antigen (PCNA).	Daucus carota	37,914	30-Sep-99

Table 4 (continued)

	GB_PL2:AC006267	101644	AC006267	Arabidopsis thaliana BAC F9M13 from chromosome IV near 21.5 cM, Arabidopsis thaliana complete sequence.	36,158	27-Apr-99
rx00471 813	GB_BA1:TT10SARNA	721	Y15063	Thermus thermophilus 10Sa RNA gene.	39,494	18-Aug-98
	GB_BA1:SERERYAA	11219	M63676	S. erythraea first ORF of eryA gene, complete cds.	38,781	26-Apr-93
	GB_PAT:AR049367	11219	AR049367	Sequence 1 from patent US 5824513.	38,781	29-Sep-99
	GB_BA1:SERERYAA	11219	M63676	S. erythraea first ORF of eryA gene, complete cds.	38,205	26-Apr-93
rx00499 1404	GB_PR4:AC007206	42732	AC007206	Homo sapiens chromosome 19, cosmid R27370, complete sequence. Homo sapiens	34,982	4-Apr-99
	GB_EST26:AI344735	462	AI344735	qp05a10.x1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1917114 3' similar to gb:IM15800 T-LYMPHOCYTE MATURATION-ASSOCIATED PROTEIN (HUMAN);, mRNA sequence.	42,675	2-Feb-99
rx00500 798	GB_PR4:AC006479	161837	AC006479	Homo sapiens clone DJ1051J04, complete sequence.	38,462	11-Nov-99
	GB_PR4:AC006111	190825	AC006111	Homo sapiens chromosome 16 clone RPC1-11_461A8, complete sequence.	40,736	3-Jul-99
	GB_HTG2:AF128834	196589	AF128834	Homo sapiens chromosome 8 clone BAC 57G24 map 8p12, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	34,062	28-Feb-99
	GB_HTG2:AF128834	196589	AF128834	Homo sapiens chromosome 8 clone BAC 57G24 map 8p12, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	34,062	28-Feb-99
rx00501 630	GB_BA1:D86429	5925	D86429	Saccharopolyspora rectivirgula gene for beta-galactosidase, complete cds.	53,871	09-DEC-1998
	GB_HTG1:HS1099D15	1301	AL035456	Homo sapiens chromosome 20 clone RP5-1099D15, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	33,546	23-Nov-99
	GB_HTG1:HS1099D15	1301	AL035456	Homo sapiens chromosome 20 clone RP5-1099D15, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	33,546	23-Nov-99
rx00502 1155	GB_BA2:U00015	42325	U00015	Mycobacterium leprae cosmid B1620.	34,783	01-MAR-1994
	GB_BA1:U00020	36947	U00020	Mycobacterium leprae cosmid B229.	34,900	01-MAR-1994
	GB_HTG1:HS179115	210672	Z84464	Homo sapiens chromosome 13 clone 179115, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	32,898	22-Jan-97
	GB_BA1:MTV008	63033	AL021246	Mycobacterium tuberculosis H37Rv complete genome; segment 108/162.	37,011	17-Jun-98
rx00566 729	GB_BA2:AF071885	2188	AF071885	Streptomyces coelicolor ATP-dependent Clp protease proteolytic subunit 1 (clpP1) and ATP-dependent Clp protease proteolytic subunit 2 (clpP2) genes, complete cds; and ATP-dependent Clp protease ATP-binding subunit ClpX (clpX) gene, partial cds.	62,963	28-Jun-99
	GB_BA2:AF013216	15742	AF013216	Myxococcus xanthus Dog (dog), isocitrate lyase (icl), Mls (mls), Ufo (ufo), fumarate hydratase (fhy), and proteosome major subunit (clpP) genes, complete cds; and acyl-CoA oxidase (aco) gene, partial cds.	54,683	28-Jan-98

Table 4 (continued)

rx00567	714	GB_BA1:MTV008	63033	AL021246	Mycobacterium tuberculosis H37Rv complete genome; segment 108/162.	Mycobacterium tuberculosis	42,090	17-Jun-98
		GB_BA1:CGBPHI16	962	Y12472	C. glutamicum DNA, attachment site bacteriophage Phi-16.	Corynebacterium glutamicum	40,000	05-MAR-1999
		GB_BA1:ECOCPLPPA	1236	J05534	Escherichia coli ATP-dependent dlp protease proteolytic component (clpP) gene, complete cds.	Escherichia coli	52,119	26-Apr-93
rx00621	906	GB_EST1:D36491	360	D36491	CELK033GYF Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA clone yk33g11 5', mRNA sequence.	Caenorhabditis elegans	40,390	8-Aug-94
		GB_IN2:CELC16A3	34968	U41534	Caenorhabditis elegans cosmid C16A3.	Caenorhabditis elegans	35,477	18-MAY-1999
		GB_HTG3:AC009311	160198	AC009311	Homo sapiens clone NH0311L03. *** SEQUENCING IN PROGRESS	Homo sapiens	38,636	13-Aug-99
					*** 3 unordered pieces.			
rx00622	1539	GB_BA1:AB004795	3039	AB004795	Pseudomonas sp. gene for dipeptidyl aminopeptidase, complete cds.	Pseudomonas sp.	54,721	5-Feb-99
		GB_BA1:MBOP11	2392	D38405	Moraxella lacunata gene for protease II, complete cds.	Moraxella lacunata	50,167	8-Feb-99
		GB_IN2:AF078916	2960	AF078916	Trypanosoma brucei brucei oligopeptidase B (opb) gene, complete cds.	Trypanosoma brucei	48,076	08-OCT-1999
rx00650	759	GB_BA2:AF161327	2021	AF161327	Corynebacterium diphtheriae histidine kinase ChrS (chrS) and response regulator ChrA (chrA) genes, complete cds.	Corynebacterium diphtheriae	51,319	9-Sep-99
		GB_PL2:ATAC006533	99188	AC006533	Arabidopsis thaliana chromosome II BAC F20M17 genomic sequence, complete cds.	Arabidopsis thaliana	38,051	26-MAY-1999
		GB_PL2:ATAC006533	99188	AC006533	Arabidopsis thaliana chromosome II BAC F20M17 genomic sequence, complete cds.	Arabidopsis thaliana	35,403	26-MAY-1999
rx00675	915	GB_BA1:SC3C8	33095	AL023861	Streptomyces coelicolor cosmid 3C8.	Streptomyces coelicolor	36,836	15-Jan-99
		GB_PR3:AC005736	215441	AC005736	Homo sapiens chromosome 16, BAC clone 462G18 (LANL), complete sequence.	Homo sapiens	42,027	01-OCT-1998
		GB_IN2:AC005719	188357	AC005719	Drosophila melanogaster, chromosome 2L, region 38A5-38B4, BAC clone BACR48M05, complete sequence.	Drosophila melanogaster	35,531	27-OCT-1999
rx00689	1614	GB_PAT:E07294	2975	E07294	genomic DNA encoding dehydrogenase of Bacillus stearothermophilus.	Bacillus stearothermophilus	45,677	29-Sep-97
		GB_BA1:BACALDHT	1975	D13846	B. stearothermophilus aldH gene for aldehyde dehydrogenase, complete cds.	Bacillus stearothermophilus	45,677	20-Feb-99
		GB_BA2:PPU96338	5276	U96338	Pseudomonas putida NCIMB 9866 plasmid pRA4000 p-cresol degradative pathway genes, p-hydroxybenzaldehyde dehydrogenase (pchA), p-cresol methylhydroxylase, cytochrome subunit precursor (pchC), unknown (pchX) and p-cresol methylhydroxylase, flavoprotein subunit (pchF) genes, complete cds.	Pseudomonas putida	44,317	13-MAY-1999
rx00715	918	GB_EST30:A1647104	218	A1647104	vn15c01.y1 Stratagene mouse heart (#937316) Mus musculus cDNA clone IMAGE:1021248 5', mRNA sequence.	Mus musculus	58,511	29-Apr-99
		GB_EST17:AA636159	447	AA636159	vn15c01.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone IMAGE:1021248 5', mRNA sequence.	Mus musculus	41,195	22-OCT-1997

Table 4 (continued)

GB_EST10:AA184468	583	AA184468	mi52h05.r1 Stratagene mouse embryonic carcinoma (#937317) Mus musculus cDNA clone IMAGE:633561 5' similar to gb:D10918 Mouse mRNA for ubiquitin like protein, partial sequence (MOUSE);, mRNA sequence.	40,426	12-Feb-97
GB_HTG3:AC009855	167592	AC009855	Homo sapiens clone 1_C_5, *** SEQUENCING IN PROGRESS ***	36,673	3-Sep-99
GB_HTG3:AC009855	167592	AC009855	13 unordered pieces.	36,673	3-Sep-99
GB_PR4:AC005082	169739	AC005082	Homo sapiens clone RG271G13, complete sequence.	39,557	8-Sep-99
GB_BA1:MLCB596	38426	AL035472	Mycobacterium leprae cosmid B596.	54,562	27-Aug-99
GB_GSS12:AC0368028	652	AQ368028	tox0001N11r CUGI Tomato BAC Library Lycopersicon esculentum genomic clone tox0001N11r, genomic survey sequence.	42,557	5-Feb-99
GB_HTG3:AC008067	151242	AC008067	Homo sapiens clone NH0303104, *** SEQUENCING IN PROGRESS ***	37,239	8-Sep-99
GB_BA1:MLU15182	40123	U15182	*** 2 unordered pieces.	36,616	09-MAR-1995
GB_BA1:MSGL611CS	37769	L78822	Mycobacterium leprae cosmid L611 DNA sequence.	35,714	15-Jun-96
GB_GSS14:AC0578181	728	AQ0578181	nbxb0083P08r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0083P08r, genomic survey sequence.	39,246	2-Jun-99
GB_GSS5:AC0769737	519	AQ769737	HS_3160_A2_G04_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3160 Col=8 Row=M, genomic survey sequence.	37,765	28-Jul-99
GB_BA1:RTU08434	2400	U08434	Rhizobium trifolii orotate phosphoribosyltransferase (pyrE) and fructokinase (frk) genes, complete cds.	40,700	16-Apr-97
GB_EST31:F33810	243	F33810	HSPD27491 HM3 Homo sapiens cDNA clone s3000041E12, mRNA sequence.	41,564	13-MAY-1999
GB_PR4:AC005868	96180	AC005868	Homo sapiens 12q24.2 PAC RPCI5-944M2 (Roswell Park Cancer Institute Human PAC Library) complete sequence.	32,298	27-Feb-99
GB_EST8:AA000903	396	AA000903	mg38b04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:426031 5', mRNA sequence.	42,045	18-Jul-96
GB_EST25:AI317789	696	AI317789	uj20g09.y1 Sugano mouse embryo mewa Mus musculus cDNA clone IMAGE:1920544 5' similar to WP:C13C4.5 CE08130 SUGAR TRANSPORTER, mRNA sequence.	38,557	17-DEC-1998
GB_PH:BPH6589	41489	AJ06589	Bacteriophage phi-C31 complete genome.	41,806	29-Apr-99
GB_HTG2:AC006887	215801	AC006887	Caenorhabditis elegans clone Y59H11, *** SEQUENCING IN PROGRESS ***	35,798	24-Feb-99
GB_HTG2:AC006887	215801	AC006887	Caenorhabditis elegans clone Y59H11, *** SEQUENCING IN PROGRESS ***	35,798	24-Feb-99
GB_GSS15:AC0605195	459	AQ605195	HS_2136_B1_C12_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2136 Col=23 Row=F, genomic survey sequence.	38,074	10-Jun-99

Table 4 (continued)

GB_HTG1:CNS00M8S	214599	AL079302	Homo sapiens chromosome 14 clone R-1089B7, *** SEQUENCING IN PROGRESS ***; in ordered pieces.	Homo sapiens	38,120	15-OCT-1999
GB_HTG1:CNS00M8S	214599	AL079302	Homo sapiens chromosome 14 clone R-1089B7, *** SEQUENCING IN PROGRESS ***; in ordered pieces.	Homo sapiens	38,120	15-OCT-1999
GB_BA1:CGORF4GEN	2398	X95649	C.glutamicum ORF4 gene.	Corynebacterium glutamicum	99,273	10-MAR-1998
GB_BA1:BLDAPAB	3572	Z21502	B.lactofementum dapA and dapB genes for dihydrodipicolinate synthase and dihydrodipicolinate reductase.	Corynebacterium glutamicum	99,301	16-Aug-93
GB_PAT:E14517	1411	E14517	DNA encoding Brevibacterium dihydrodipicolinic acid reductase.	Corynebacterium glutamicum	99,659	28-Jul-99
GB_PAT:I92050	567	I92050	Sequence 17 from patent US 5726299.	Unknown.	62,787	01-DEC-1998
GB_PAT:I78760	567	I78760	Sequence 16 from patent US 5693781.	Unknown.	62,787	3-Apr-98
GB_BA2:AE000426	10240	AE000426	Escherichia coli K-12 MG1655 section 316 of 400 of the complete genome.	Escherichia coli	36,456	12-Nov-98
GB_BA2:AE001598	11136	AE001598	Chlamydia pneumoniae section 14 of 103 of the complete genome.	Chlamydia pneumoniae	32,782	08-MAR-1999
GB_PL2:AF079370	2897	AF079370	Kluyveromyces lactis invertase (INV1) gene, complete cds.	Kluyveromyces lactis	35,849	4-Aug-99
GB_BA2:AE001598	11136	AE001598	Chlamydia pneumoniae section 14 of 103 of the complete genome.	Chlamydia pneumoniae	40,138	08-MAR-1999
GB_PR2:HSQ15C24	73192	AJ239325	Homo sapiens chromosome 21 from cosmids LLNLc116 1C16 and LLNLc116 15C24 map 21q22.3 region D21S171-LA161, complete sequence.	Homo sapiens	35,076	28-Sep-99
GB_GSS4:AQ691923	446	AQ691923	HS_5400_B2_G04_SP6E RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=976 Col=8 Row=N, genomic survey sequence.	Homo sapiens	33,500	6-Jul-98
GB_EST37:A1967802	479	A1967802	Ljirnp12-930-d6 Ljirnp Lambda HybriZap two-hybrid library Lotus japonicus cDNA clone LP930-12-d6 5' similar to 60S ribosomal protein L7A, mRNA sequence.	Lotus japonicus	41,127	24-Aug-99
GB_PAT:I78750	588	I78750	Sequence 6 from patent US 5693781.	Unknown.	97,071	3-Apr-98
GB_PAT:I92039	588	I92039	Sequence 6 from patent US 5726299.	Unknown.	97,071	01-DEC-1998
GB_PR3:HS929C8	139190	AL020994	Human DNA sequence from clone 929C8 on chromosome 22q12.1-12.3 Contains CA repeat, GSS, STS, complete sequence.	Homo sapiens	39,016	23-Nov-99
GB_PAT:I78750	588	I78750	Sequence 6 from patent US 5693781.	Unknown.	97,561	3-Apr-98
GB_PAT:I92039	588	I92039	Sequence 6 from patent US 5726299.	Unknown.	97,561	01-DEC-1998
GB_PAT:I78750	588	I78750	Sequence 6 from patent US 5693781.	Unknown.	37,222	3-Apr-98
GB_BA1:AB032799	9077	AB032799	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B, vio C, vio D), complete cds.	Chromobacterium violaceum	39,868	02-OCT-1999
GB_BA2:AF172851	10094	AF172851	Chromobacterium violaceum violacein biosynthetic gene cluster, complete sequence.	Chromobacterium violaceum	42,760	30-Aug-99
GB_BA1:AB032799	9077	AB032799	Chromobacterium violaceum violacein biosynthetic gene cluster (vio A, vio B, vio C, vio D), complete cds.	Chromobacterium violaceum	39,551	02-OCT-1999

Table 4 (continued)

rx000982	1629	GB_BA1:BLARGS	2501	Z21501	B.lactofermentum argS and lysA genes for arginyl-tRNA synthetase and diaminopimelate decarboxylase (partial).	Corynebacterium glutamicum	39,003	28-DEC-1993
		GB_BA1:CGXLYSA	2344	X54740	Corynebacterium glutamicum argS-lysA operon gene for the upstream region of the arginyl-tRNA synthetase and diaminopimelate decarboxylase (EC 4.1.1.20).	Corynebacterium glutamicum	41,435	30-Jun-93
rx000983	1599	GB_PAT:E14508	3579	E14508	DNA encoding Brevibacterium diaminopimelic acid decarboxylase and arginyl-tRNA synthetase.	Corynebacterium glutamicum	40,566	28-Jul-99
		GB_HTG2:AC008152	24000	AC008152	Leishmania major chromosome 35 clone L7936 strain Friedlin, *** SEQUENCING IN PROGRESS ***; 4 unordered pieces.	Leishmania major	38,658	28-Jul-99
		GB_HTG2:AC008152	24000	AC008152	Leishmania major chromosome 35 clone L7936 strain Friedlin, *** SEQUENCING IN PROGRESS ***; 4 unordered pieces.	Leishmania major	38,658	28-Jul-99
		GB_HTG3:AC008648	87249	AC008648	Homo sapiens chromosome 5 clone CIT978SKB_186E14, *** SEQUENCING IN PROGRESS ***; 22 unordered pieces.	Homo sapiens	36,102	3-Aug-99
rx000984	440	GB_BA1:MVINED	3098	D01045	Micromonospora viridifaciens DNA for nadR protein and neuraminidase, complete cds.	Micromonospora viridifaciens	59,226	2-Feb-99
		GB_PAT:E02375	1881	E02375	Neuraminidase gene.	Micromonospora viridifaciens	59,226	29-Sep-97
		GB_PR4:HUAC004513	101311	AC004513	Homo sapiens Chromosome 16 BAC clone CIT987SKA-926E7, complete sequence.	Homo sapiens	41,204	23-Nov-99
rx001014	2724	GB_BA1:MTV008	63033	AL021246	Mycobacterium tuberculosis H37Rv complete genome; segment 108/162.	Mycobacterium tuberculosis	56,167	17-Jun-98
		GB_BA1:STMAMPEPN	2849	L23172	Streptomyces lividans aminopeptidase N gene, complete cds.	Streptomyces lividans	57,067	18-MAY-1994
		GB_BA1:SC7H2	42655	AL109732	Streptomyces coelicolor cosmid 7H2.	Streptomyces coelicolor A3(2)	37,551	2-Aug-99
rx001059	732	GB_HTG3:AC008154	172241	AC008154	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS ***; 26 unordered pieces.	Homo sapiens	39,499	8-Sep-99
		GB_HTG3:AC008154	172241	AC008154	Homo sapiens chromosome 7, *** SEQUENCING IN PROGRESS ***; 26 unordered pieces.	Homo sapiens	39,499	8-Sep-99
		GB_EST32:AI756574	299	AI756574	ead2f10.y1 Eimeria M5-6 Merozoite stage Eimeria tenella cDNA 5', mRNA sequence.	Eimeria tenella	37,793	23-Jun-99
rx001073	954	GB_BA1:BACOUTB	1004	M15811	Bacillus subtilis outB gene encoding a sporulation protein, complete cds.	Bacillus subtilis	53,723	26-Apr-93
		GB_PR4:AC007938	167237	AC007938	Homo sapiens clone UWGC:djs201 from 7q31, complete sequence.	Homo sapiens	34,322	1-Jul-99
		GB_PL2:ATAC006282	92577	AC006282	Arabidopsis thaliana chromosome II BAC F13K3 genomic sequence, complete sequence.	Arabidopsis thaliana	36,181	13-MAR-1999
rx001120	1401	GB_BA1:MTV008	63033	AL021246	Mycobacterium tuberculosis H37Rv complete genome; segment 108/162.	Mycobacterium tuberculosis	36,715	17-Jun-98
		GB_BA1:CAJ10321	6710	AJ010321	Caulobacter crescentus partial tig gene and clpP, clpX, lon genes.	Caulobacter crescentus	63,311	01-OCT-1998

Table 4 (continued)

GB_BA2:AF150957	4440	AF150957	Azospirillum brasilense trigger factor (tig), heat-shock protein ClpP (clpP), and heat-shock protein ClpX (clpX) genes, complete cds; and Lon protease (lon) gene, partial cds.	Azospirillum brasilense	60,813	7-Jun-99
GB_PR3:HS408N23	97916	Z98048	Human DNA sequence from PAC 408N23 on chromosome 22q13. Contains HIP, HSC70-INTERACTING PROTEIN (PROGESTERONE RECEPTOR-ASSOCIATED P48 PROTEIN), ESTs and STS.	Homo sapiens	34,567	23-Nov-99
GB_BA2:AE001227	26849	AE001227	Treponema pallidum section 43 of 87 of the complete genome.	Treponema pallidum	37,564	16-Jul-98
GB_PR3:HS408N23	97916	Z98048	Human DNA sequence from PAC 408N23 on chromosome 22q13. Contains HIP, HSC70-INTERACTING PROTEIN (PROGESTERONE RECEPTOR-ASSOCIATED P48 PROTEIN), ESTs and STS.	Homo sapiens	34,911	23-Nov-99
GB_BA1:MTCY261	27322	Z97559	Mycobacterium tuberculosis H37Rv complete genome; segment 95/162.	Mycobacterium tuberculosis	38,789	17-Jun-98
GB_HTG4:AC009849	114993	AC009849	Drosophila melanogaster chromosome 2 clone BACR07H08 (D864) RPCI-98 07.H.8 map 31B-31C strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 55 unordered pieces.	Drosophila melanogaster	39,213	25-OCT-1999
GB_HTG4:AC009849	114993	AC009849	Drosophila melanogaster chromosome 2 clone BACR07H08 (D864) RPCI-98 07.H.8 map 31B-31C strain Y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 55 unordered pieces.	Drosophila melanogaster	39,213	25-OCT-1999
GB_BA2:AF176799	2943	AF176799	Lactobacillus pentosus PapQ (papQ) and catabolite control protein A (ccpA) genes, complete cds.	Lactobacillus pentosus	37,043	5-Sep-99
GB_BA2:AF012084	3082	AF012084	Lactobacillus helveticus prolidase (papQ) gene, complete cds.	Lactobacillus helveticus	46,796	1-Jul-98
GB_EST32:AT728955	611	AT728955	BNLGH112114 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (AC004481) putative permease (Arabidopsis thaliana), mRNA sequence.	Gossypium hirsutum	37,647	11-Jun-99
GB_BA1:MLCB22	40281	Z98741	Mycobacterium leprae cosmid B22.	Mycobacterium leprae	61,570	22-Aug-97
GB_BA1:MTCY190	34150	Z70283	Mycobacterium tuberculosis H37Rv complete genome; segment 98/162.	Mycobacterium tuberculosis	60,434	17-Jun-98
GB_BA1:SC5F7	40024	AL098872	Streptomyces coelicolor cosmid 5F7.	Streptomyces coelicolor A3(2)	57,011	22-Jul-99
GB_HTG1:CEY116A8	2110000	Z98858	Caenorhabditis elegans chromosome IV clone Y116A8, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	34,843	26-Oct-99
GB_HTG1:CEY116A8	2110000	Z98858	Caenorhabditis elegans chromosome IV clone Y116A8, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	34,843	26-Oct-99
GB_IN1:CEY116A8C	260341	AL117204	Caenorhabditis elegans cosmid Y116A8C, complete sequence.	Caenorhabditis elegans	34,843	19-Nov-99
GB_BA1:D90915	130001	D90915	Synechocystis sp. PCC6803 complete genome, 17127, 2137259-2267259.	Synechocystis sp.	36,538	7-Feb-99
GB_BA1:D90915	130001	D90915	Synechocystis sp. PCC6803 complete genome, 17127, 2137259-2267259.	Synechocystis sp.	34,512	7-Feb-99



Table 4 (continued)

GB_HTG3:AC010515	41038	AC010515	Homo sapiens chromosome 19 clone LLNL-R_249H9, *** SEQUENCING IN PROGRESS ***; 31 unordered pieces.	33,564	15-Sep-99
GB_OM:CFP180RRC	5425	X87224	Canis familiaris mRNA for ribosome receptor, p180.	41,229	22-Jan-99
GB_OM:CFP180RRC	5425	X87224	Canis familiaris mRNA for ribosome receptor, p180.	38,167	22-Jan-99
GB_IN1:CEY47D3A	199814	AL117202	Caenorhabditis elegans cosmid Y47D3A, complete sequence.	36,604	19-Nov-99
GB_PR4:AC006039	176257	AC006039	Homo sapiens clone NH0319F03, complete sequence.	34,984	05-MAY-1999
GB_PR4:AC006039	176257	AC006039	Homo sapiens clone NH0319F03, complete sequence.	35,951	05-MAY-1999
GB_EST22:A070047	479	A070047	UI-R-C1-in-f08-0-UI.s1 UI-R-C1 Rattus norvegicus cDNA clone UI-R; Rattus norvegicus C1-in-f08-0-UI 3', mRNA sequence.	36,975	5-Jul-99
GB_RO:S75965	625	S75965	THP=Tamm-Horsfall protein [promoter] [rats, Genomic, 625 nt].	34,400	27-Jul-95
GB_EST5:H96951	459	H96951	yu01g03.r1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone IMAGE:232564 5', mRNA sequence.	32,969	11-DEC-1995
GB_PL1:NEULCCB	2656	M18334	N.crassa (strain TS) lacase gene, complete cds.	44,330	03-MAY-1994
GB_OV:WTRACOMPL	16714	Y16884	Rhea americana complete mitochondrial genome.	35,094	19-Jul-99
GB_OV:AF090339	16704	AF090339	Rhea americana mitochondrion, complete genome.	35,094	27-MAY-1999
GB_PL2:AF111709	52684	AF111709	Oryza sativa subsp. indica Retrosat 1 retrotransposon and Ty3-Gypsy type Retrosat 2 retrotransposon, complete sequences; and unknown genes.	37,410	26-Apr-99
GB_IN1:CELZC250	34372	AF003383	Caenorhabditis elegans cosmid ZC250.	35,506	14-MAY-1997
GB_EST1:Z14808	331	Z14808	CEL5E4 Chris Martin sorted cDNA library Caenorhabditis elegans cDNA clone cm5e4 5', mRNA sequence.	36,890	19-Jun-97
GB_BA1:MTIC65	34331	Z95584	Mycobacterium tuberculosis H37Rv complete genome; segment 50/162.	59,298	17-Jun-98
GB_BA1:MSGY348	40056	AD000020	Mycobacterium tuberculosis sequence from clone y348.	59,227	10-DEC-1996
GB_BA1:SC5C7	41906	AL031515	Streptomyces coelicolor cosmid 5C7.	39,261	7-Sep-98
GB_BA1:TTAJ5043	837	AJ225043	Thermus thermophilus partial narK gene.	55,245	18-Jun-98
GB_PL2:AC010675	84723	AC010675	Arabidopsis thaliana chromosome 1 BAC T17F3 genomic sequence, complete sequence.	37,058	11-Nov-99
GB_GSS9:AQ170862	518	AQ170862	HS_3165_B2_F03_T7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3165 Col=6 Row=L, genomic survey sequence.	38,610	17-OCT-1998
GB_BA1:D90757	17621	D90757	Escherichia coli genomic DNA. (27.3 - 27.7 min).	55,445	7-Feb-99
GB_BA1:D90787	15942	D90787	E.coli genomic DNA, Kohara clone #276(33.0-33.3 min.).	36,815	29-MAY-1997
GB_BA1:D90758	13660	D90758	Escherichia coli genomic DNA. (27.6 - 27.9 min).	54,942	7-Feb-99
GB_BA1:SCJ12	35302	AL109989	Streptomyces coelicolor cosmid J12.	62,423	24-Aug-99
GB_BA1:BSNARYWI	12450	Z49884	B.subtilis narf(G.H.I.J.K), yw[C.D.E] and argS genes.	57,447	24-Jun-98

Table 4 (continued)

rx01358	1644	GB_BA1:BSUB0020	212150	Z99123	Bacillus subtilis complete genome (section 20 of 21): from 3798401 to Bacillus subtilis 4010550.	37,129	26-Nov-97
		GB_GSS1:AAQ260413	453	AQ260413	CITBI-E1-2510B12.TF CITBI-E1 Homo sapiens genomic clone	41,531	24-OCT-1998
		GB_EST20:AA840582	326	AA840582	2510B12, genomic survey sequence.	42,901	27-Feb-98
					vw77h07.r1 Stratagene mouse heart (#937316) Mus musculus cDNA Mus musculus clone IMAGE:1261021 5' similar to gb.J04181 Mouse A-X actin mRNA, complete cds (MOUSE);, mRNA sequence.		
		GB_PAT:A39944	3836	A39944	Sequence 1 from Patent WO9421807.	38,764	05-MAR-1997
rx01385	2004	GB_BA1:FVBPENTA	2519	M98557	Flavobacterium sp. pentachlorophenol 4-monoxygenase gene, complete mRNA.	40,855	26-Apr-93
		GB_PAT:I19994	2516	I19994	Sequence 2 from patent US 5512478.	40,855	07-OCT-1996
		GB_BA2:AF059680	2410	AF059680	Sphingomonas sp. UG30 pentachlorophenol 4-monoxygenase (pcpB) gene, complete cds; and pentachlorophenol 4-monoxygenase reductase (pcpD) gene, partial cds.	42,993	27-Apr-99
rx01412	327	GB_GSS12:AAQ332469	459	AQ332469	HS_5003_A1_H08_SP6E RPC11 Human Male BAC Library Homo sapiens genomic clone Plate=579 Col=15 Row=O, genomic survey sequence.	38,208	05-MAR-1999
		GB_EST27:AA998532	453	AA998532	UI-R-C0-ic-d-11-0-UI.s1 UI-R-C0 Rattus norvegicus cDNA clone UI-R-C0-ic-d-11-0-UI 3', mRNA sequence.	39,336	09-MAR-1999
		GB_HTG1:HA342D11	178183	AL121748	Homo sapiens chromosome 10 clone RP11-342D11, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	40,550	23-Nov-99
rx01458	1173	GB_BA2:AE000745	15085	AE000745	Aquifex aeolicus section 77 of 109 of the complete genome.	37,694	25-MAR-1998
		GB_BA2:AE000745	15085	AE000745	Aquifex aeolicus section 77 of 109 of the complete genome.	35,567	25-MAR-1998
rx01571	723	GB_BA1:AB011413	12070	AB011413	Streptomyces griseus genes for Orf2, Orf3, Orf4, Orf5, AfsA, Orf8, partial and complete cds.	57,500	7-Aug-98
		GB_BA1:AB011413	12070	AB011413	Streptomyces griseus genes for Orf2, Orf3, Orf4, Orf5, AfsA, Orf8, partial and complete cds.	35,655	7-Aug-98
rx01607	753	GB_PR4:AC005005	133893	AC005005	Homo sapiens PAC clone DJ412A9 from 22, complete sequence.	38,399	02-MAR-1999
		GB_HTG3:AC008257	109187	AC008257	Drosophila melanogaster chromosome 2 clone BACR08A11 (D916) RPC1-98.08.A.11 map 42A-42A strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 93 unordered pieces.	33,741	08-OCT-1999
		GB_HTG3:AC008257	109187	AC008257	Drosophila melanogaster chromosome 2 clone BACR08A11 (D916) RPC1-98.08.A.11 map 42A-42A strain y; cn bw sp. *** SEQUENCING IN PROGRESS ***; 93 unordered pieces.	33,741	08-OCT-1999
rx01609	996	GB_BA1:MTV003	13246	AL008883	Mycobacterium tuberculosis H37Rv complete genome; segment 125/162.	39,369	17-Jun-98
		GB_BA1:MSG81529CS	36985	L78824	Mycobacterium leprae cosmid B1529 DNA sequence.	60,624	15-Jun-96
		GB_BA1:AB024601	14807	AB024601	Pseudomonas aeruginosa dapD gene for tetrahydrodipicolinate N-succinyltransferase, complete cds, strain PAO1.	41,603	12-MAR-1999

Table 4 (continued)

rx01654	1119	GB_GSS4:AQ704352	532	AQ704352	HS_2147_A2_H04_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2147 Col=8 Row=O, genomic survey sequence.	Homo sapiens	37,838	7-Jul-99
		GB_RO:MMAE000663	250611	AE000663	Mus musculus TCR beta locus from bases 1 to 250611 (section 1 of 3) of the complete sequence.	Mus musculus	35,799	4-Sep-97
		GB_EST23:AI158428	511	AI158428	IMAGE:1446863 5', mRNA sequence.	Mus musculus	41,337	30-Sep-98
rx01664	945	GB_OV:AF026198	63155	AF026198	Fugu rubripes neural cell adhesion molecule L1 homolog (L1-CAM) gene, complete cds; putative protein 1 (PUT1) gene, partial cds; mitosis-specific chromosome segregation protein SMC1 homolog (SMC1) gene, complete cds; and calcium channel alpha-1 subunit homolog (CCA1) and putative protein 2 (PUT2) genes, partial cds, complete sequence.	Fugu rubripes	35,187	02-MAY-1998
		GB_PR3:AC004466	122186	AC004466	Homo sapiens 12q13.1 PAC RPC15-1057120 (Roswell Park Cancer Institute Human PAC library) complete sequence.	Homo sapiens	37,382	17-Sep-98
		GB_PR3:AC004466	122186	AC004466	Homo sapiens 12q13.1 PAC RPC15-1057120 (Roswell Park Cancer Institute Human PAC library) complete sequence.	Homo sapiens	37,325	17-Sep-98
rx01795	720	GB_BA2:CGU13922	4412	U13922	Corynebacterium glutamicum putative type II 5-cytosine methyltransferase (cgIM) and putative type II restriction endonuclease (cgIR) and putative type I or type III restriction endonuclease (cgIR) genes, complete cds.	Corynebacterium glutamicum	99,444	3-Feb-98
		GB_BA1:S86113	1044	S86113	ORF 1 [Neisseria gonorrhoeae, Genomic, 1044 nt].	Neisseria gonorrhoeae	58,320	07-MAY-1993
		GB_PAT:122080	850	122080	Sequence 1 from patent US 5525717.	Unknown.	57,722	07-OCT-1996
rx01802	954	GB_BA2:AE001519	14062	AE001519	Helicobacter pylori, strain J99 section 80 of 132 of the complete genome.	Helicobacter pylori J99	33,510	20-Jan-99
		GB_GSS5:AQ774071	552	AQ774071	HS_2269_B1_C10_TTC CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2269 Col=19 Row=F, genomic survey sequence.	Homo sapiens	37,967	29-Jul-99
		GB_PR4:AC007459	40907	AC007459	Homo sapiens chromosome 16 clone 306C6, complete sequence.	Homo sapiens	39,140	04-MAY-1999
rx01838	842	GB_BA1:SCE15	26440	AL049707	Streptomyces coelicolor cosmid E15.	Streptomyces coelicolor	36,297	22-Apr-99
		GB_HTG3:AC009545	165042	AC009545	Homo sapiens chromosome 11 clone 131_J_04 map 11, *** SEQUENCING IN PROGRESS ***; 8 unordered pieces.	Homo sapiens	37,651	01-OCT-1999
		GB_HTG3:AC009545	165042	AC009545	Homo sapiens chromosome 11 clone 131_J_04 map 11, *** SEQUENCING IN PROGRESS ***; 8 unordered pieces.	Homo sapiens	37,651	01-OCT-1999
rx01848	867	GB_BA1:MTCY24A1	20270	Z95207	Mycobacterium tuberculosis H37Rv complete genome; segment 124/162.	Mycobacterium tuberculosis	38,270	17-Jun-98
		GB_EST21:C89252	587	C89252	C89252 Mouse early blastocyst cDNA Mus musculus cDNA clone 01B00061JC08, mRNA sequence.	Mus musculus	37,219	28-MAY-1998
		GB_EST14:AA423340	457	AA423340	ve39d04.r1 Soares mouse mammary gland N0MMG Mus musculus cDNA clone IMAGE:820519 5', mRNA sequence.	Mus musculus	38,377	16-OCT-1997

Table 4 (continued)

rx01849	1224	GB_BA1:MTCY24A1	20270	Z95207	Mycobacterium tuberculosis H37Rv complete genome; segment 124/162.	Mycobacterium tuberculosis	39,950	17-Jun-98
		GB_BA2:RCPHSYNG	45959	Z11165	R. capsulatus complete photosynthesis gene cluster.	Rhodobacter capsulatus	37,344	2-Sep-99
		GB_BA1:RSP010302	40707	AJ010302	Rhodobacter sphaeroides photosynthetic gene cluster.	Rhodobacter sphaeroides	40,898	27-Aug-99
rx01868	2049	GB_BA1:MTV033	21620	AL021928	Mycobacterium tuberculosis H37Rv complete genome; segment 11/162.	Mycobacterium tuberculosis	38,679	17-Jun-98
		GB_BA1:MLCL622	42498	Z95398	Mycobacterium leprae cosmid L622.	Mycobacterium leprae	38,911	24-Jun-97
		GB_BA1:MSGB983CS	36788	L78828	Mycobacterium leprae cosmid B983 DNA sequence.	Mycobacterium leprae	38,933	15-Jun-96
rx01885	924	GB_BA1:MTCY1A10	25949	Z95387	Mycobacterium tuberculosis H37Rv complete genome; segment 117/162.	Mycobacterium tuberculosis	51,094	17-Jun-98
		GB_PR3:HSU220B11	41247	Z69908	Human DNA sequence from cosmid cU220B11, between markers DXS6791 and DXS8038 on chromosome X.	Homo sapiens	39,038	23-Nov-99
		GB_BA1:PDU17435	993	U17435	Paracoccus denitrificans Fnr-like transcriptional activator (nnr) gene, complete cds.	Paracoccus denitrificans	39,390	19-Jul-95
rx01914	526	GB_PR3:AC005796	43843	AC005796	Homo sapiens chromosome 19, cosmid R31408, complete sequence.	Homo sapiens	34,961	06-OCT-1998
		GB_PR3:HS390C10	114231	AL008721	Homo sapiens DNA sequence from BAC 390C10 on chromosome 22q11.21-12.1. Contains an Immunoglobulin LIKE gene and a pseudogene similar to Beta Crystallin. Contains ESTs, STSs, GSSs and taga and tat repeat polymorphisms, complete sequence.	Homo sapiens	39,600	23-Nov-99
		GB_PR3:AC005796	43843	AC005796	Homo sapiens chromosome 19, cosmid R31408, complete sequence.	Homo sapiens	37,725	06-OCT-1998
rx01932	1020	GB_PR3:AC003025	112309	AC003025	Human Chromosome 11p12.2 PAC clone pJ466a11, complete sequence.	Homo sapiens	35,585	23-Jul-98
		GB_GSS3:B78728	312	B78728	CIT-HSP-431E3.TV CIT-HSP Homo sapiens genomic clone 431E3, genomic survey sequence.	Homo sapiens	38,907	25-Jun-98
		GB_PR3:AC003025	112309	AC003025	Human Chromosome 11p12.2 PAC clone pJ466a11, complete sequence.	Homo sapiens	35,859	23-Jul-98
rx01933	726	GB_HTG1:HS74O16	169401	AL110119	Homo sapiens chromosome 21 clone RPCIP704O1674 map 21q21, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	35,302	27-Aug-99
		GB_HTG1:HS74O16	169401	AL110119	Homo sapiens chromosome 21 clone RPCIP704O1674 map 21q21, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Homo sapiens	35,302	27-Aug-99
		GB_PR4:AC006032	170282	AC006032	Homo sapiens BAC clone NH0115E20 from Y, complete sequence.	Homo sapiens	37,640	27-Feb-99
rx01971	954	GB_HTG3:AC008230	108469	AC008230	Drosophila melanogaster chromosome 2 clone BACR17117 (D934) RPCI-98 17.1.17 map 53A-53C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 108 unordered pieces.	Drosophila melanogaster	35,466	10-Aug-99
		GB_HTG3:AC008230	108469	AC008230	Drosophila melanogaster chromosome 2 clone BACR17117 (D934) RPCI-98 17.1.17 map 53A-53C strain y; cn bw sp, *** SEQUENCING IN PROGRESS ***; 108 unordered pieces.	Drosophila melanogaster	35,466	10-Aug-99

Table 4 (continued)

GB_PR3:AF064860	165382	AF064860	Homo sapiens chromosome 21q22.3 PAC 70124, complete sequence.	Homo sapiens	39,716	2-Jun-98
GB_EST2:D48846	459	D48846	RICS15292A Rice green shoot Oryza sativa cDNA, mRNA sequence.	Oryza sativa	37,118	2-Aug-95
GB_GSS10:AQ195886	595	AQ195886	RPC111-66O13.TJ RPC1-11 Homo sapiens genomic clone RPC1-11-66O13, genomic survey sequence.	Homo sapiens	41,000	20-Apr-99
GB_GSS10:AQ195886	595	AQ195886	RPC111-66O13.TJ RPC1-11 Homo sapiens genomic clone RPC1-11-66O13, genomic survey sequence.	Homo sapiens	34,790	20-Apr-99
GB_EST20:AA855266	406	AA855266	vw70b08.r1 Striatagene mouse heart (#937316) Mus musculus cDNA clone IMAGE:1260279 5', mRNA sequence.	Mus musculus	42,638	06-MAR-1998
GB_EST20:AA855266	406	AA855266	vw70b08.r1 Striatagene mouse heart (#937316) Mus musculus cDNA clone IMAGE:1260279 5', mRNA sequence.	Mus musculus	37,183	06-MAR-1998
GB_BA1:SC5C7	41906	AL031515	Streptomyces coelicolor cosmid 5C7.	Streptomyces coelicolor	41,732	7-Sep-98
GB_BA1:MTCl65	34331	Z95584	Mycobacterium tuberculosis H37Rv complete genome; segment 50/162.	Mycobacterium tuberculosis	62,395	17-Jun-98
GB_BA1:SCJ12	35302	AL109989	Streptomyces coelicolor cosmid J12.	Streptomyces coelicolor A3(2)	61,503	24-Aug-99
GB_PAT:E15823	2323	E15823	DNA encoding cell surface protein from Corynebacterium ammoniagenes.	Corynebacterium ammoniagenes	53,942	28-Jul-99
GB_OM:SSAMPTDN	3387	Z29522	S.scrofa mRNA for aminopeptidase N.	Sus scrofa	42,672	26-Sep-94
GB_OV:D87992	3181	D87992	Gallus gallus mRNA for aminopeptidase Ey, complete cds.	Gallus gallus	41,554	5-Jun-99
GB_BA1:AP000064	247695	AP000064	Aeropyrum pernix genomic DNA, section 777.	Aeropyrum pernix	39,882	22-Jun-99
GB_PL2:ATAC006587	79262	AC006587	Arabidopsis thaliana chromosome II BAC T17D12 genomic sequence, complete sequence.	Arabidopsis thaliana	38,490	23-MAR-1999
GB_PL2:ATAC006587	79262	AC006587	Arabidopsis thaliana chromosome II BAC T17D12 genomic sequence, complete sequence.	Arabidopsis thaliana	34,863	23-MAR-1999
GB_BA2:AF120718	4137	AF120718	Lactobacillus fermentum urease operon, partial sequence.	Lactobacillus fermentum	56,265	31-MAR-1999
GB_PAT:E03531	2896	E03531	DNA sequence coding for acid urease.	Lactobacillus fermentum	56,265	29-Sep-97
GB_BA1:LBAAURE	2896	D10805	L.fermentum gene for acid urease.	Lactobacillus fermentum	56,265	2-Feb-99
GB_GSS10:AQ242920	451	AQ242920	HS_2061_A1_E08_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2061 Col=15 Row=I, genomic survey sequence.	Homo sapiens	37,916	03-OCT-1998
GB_IN1:SLMMTPMF	14503	D29637	Physarum polycephalum mitochondrial DNA.	Mitochondrion Physarum polycephalum	40,335	12-MAY-1999
GB_IN2:AF012249	5542	AF012249	Physarum polycephalum strain aux2-S region of mitochondria derived from mtF plasmid, including URFA, URFC, URFD, URFE, URFF, and URFG genes, complete cds, and URFH gene, partial cds.	Physarum polycephalum	40,335	08-MAY-1998
GB_BA2:AF048784	681	AF048784	Actinomyces naeslundii urease accessory protein (ureG) gene, complete cds.	Actinomyces naeslundii	66,814	9-Feb-99

Table 4 (continued)

GB_BA2:AF056321	5482	AF056321	Actinomyces naeslundii urease gamma subunit UreA (ureA), urease beta subunit UreB (ureB), urease alpha subunit UreC (ureC), urease accessory protein UreE (ureE), urease accessory protein UreF (ureF), urease accessory protein UreG (ureG), and urease accessory protein UreD (ureD) genes, complete cds.	63,686	9-Feb-99
GB_BA2:SSU35248	5773	U35248	Streptococcus salivarius ure cluster nickel transporter homolog (ureI) gene, partial cds, and urease beta subunit (ureA), gamma subunit (ureB), alpha subunit (ureC), and accessory proteins (ureE), (ureF), (ureG), and (ureD) genes, complete cds.	61,931	26-Jan-96
GB_GSS3:B49054	543	B49054	RPC111-4113.TV RPC1-11 Homo sapiens genomic clone RPC1-11-4113, genomic survey sequence.	39,161	8-Apr-99
GB_PL1:PMCM5GI	3363	L27092	Pneumocystis carinii B-cell receptor (msgl) gene, 3' end.	39,819	26-Sep-94
GB_PL2:AF038556	12792	AF038556	Pneumocystis carinii f. sp. hominis variant regions of major surface glycoproteins (msg1, msg3, msg4) genes, partial cds.	33,832	10-Sep-98
GB_GSS8:AQ051031	914	AQ051031	nbxb0004dG10r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0004N20r, genomic survey sequence.	32,299	24-MAR-1999
GB_GSS8:AQ051031	914	AQ051031	nbxb0004dG10r CUGI Rice BAC Library Oryza sativa genomic clone nbxb0004N20r, genomic survey sequence.	34,573	24-MAR-1999
GB_BA1:CGU35023	3195	U35023	Corynebacterium glutamicum thiosulfate sulfurtransferase (thtR) gene, partial cds, acyl CoA carboxylase (accBC) gene, complete cds.	100,000	16-Jan-97
GB_BA1:MTCY71	42729	Z92771	Mycobacterium tuberculosis H37Rv complete genome; segment 141/1162.	60,380	10-Feb-99
GB_BA1:U00012	33312	U00012	Mycobacterium leprae cosmid B1308.	37,660	30-Jan-96
GB_HTG2:HS225E12	126464	AL031772	Homo sapiens chromosome 6 clone RP1-225E12 map q24, *** SEQUENCING IN PROGRESS *** in unordered pieces.	35,973	03-DEC-1999
GB_HTG2:HS225E12	126464	AL031772	Homo sapiens chromosome 6 clone RP1-225E12 map q24, *** SEQUENCING IN PROGRESS *** in unordered pieces.	35,973	03-DEC-1999
GB_HTG2:HS225E12	126464	AL031772	Homo sapiens chromosome 6 clone RP1-225E12 map q24, *** SEQUENCING IN PROGRESS *** in unordered pieces.	36,992	03-DEC-1999
GB_BA1:AB020624	1605	AB020624	Corynebacterium glutamicum mufI gene for D-glutamate racemase, complete cds.	99,227	24-Jul-99
GB_EST4:H51527	294	H51527	yo33609.s1 Soares adult brain N2b4HB55Y Homo sapiens cDNA clone IMAGE:179705 3', mRNA sequence.	40,411	18-Sep-95
GB_GSS1:CNS003CM	1101	AL064136	Drosophila melanogaster genome survey sequence T7 end of BAC # BACR08C19 of RPC1-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.	37,674	3-Jun-99

Table 4 (continued)

rx02477	744	GB_HTG4:AC010054	130191	AC010054	Drosophila melanogaster chromosome 3L74E2 clone RPC198-15E10, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	37,466	16-OCT-1999
		GB_HTG4:AC010054	130191	AC010054	Drosophila melanogaster chromosome 3L74E2 clone RPC198-15E10, *** SEQUENCING IN PROGRESS ***; 70 unordered pieces.	Drosophila melanogaster	37,466	16-OCT-1999
		GB_HTG4:AC009375	137069	AC009375	Drosophila melanogaster chromosome 3L75A1 clone RPC198-44L18, *** SEQUENCING IN PROGRESS ***; 59 unordered pieces.	Drosophila melanogaster	39,118	16-OCT-1999
rx02513	832	GB_BA1:MTTER260	373	X92572	M.terae gene for 32 kDa protein (partial).	Mycobacterium terrae	42,895	15-Jan-98
		GB_PL1:AB019229	84294	AB019229	Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MDC16, complete sequence.	Arabidopsis thaliana	36,084	20-Nov-99
		GB_PL1:AB019229	84294	AB019229	Arabidopsis thaliana genomic DNA, chromosome 3, P1 clone: MDC16, complete sequence.	Arabidopsis thaliana	35,244	20-Nov-99
rx02531	834	GB_BA1:CGLATTB	271	X89850	C-glutamicum DNA for attB region.	Corynebacterium glutamicum	40,590	8-Aug-96
		GB_EST11:AA239557	423	AA239557	mv25f04.r1 GuayWoodford Beiler mouse kidney day 0 Mus musculus cDNA clone IMAGE:656095 5' similar to gb:X52634 Murine tm oncogene for tm protein (MOUSE); mRNA sequence.	Mus musculus	38,760	12-MAR-1997
		GB_BA1:RSPYPPCL	6500	AJ002398	Rhodobacter sphaeroides pyp and pcl genes, and orfA, orfB, orfC, orfD, orfE, orfF.	Rhodobacter sphaeroides	37,091	17-DEC-1998
rx02548	314	GB_BA2:AF127374	63734	AF127374	Streptomyces lavendulae LinA homolog, cytochrome P450 hydroxylase ORF4, cytochrome P450 hydroxylase ORF3, MitT (mitT), MitS (mitS), MitR (mitR), MitQ (mitQ), MitP (mitP), MitO (mitO), MitN (mitN), MitM (mitM), MitL (mitL), MitK (mitK), MitJ (mitJ), MitI (mitI), MitH (mitH), MitG (mitG), MitF (mitF), MitE (mitE), MitD (mitD), MitC (mitC), MitB (mitB), MitA (mitA), MmcA (mmcA), MmcB (mmcB), MmcC (mmcC), MmcD (mmcD), MmcE (mmcE), MmcF (mmcF), MmcG (mmcG), MmcH (mmch), MmcI (mmci), MmcJ (mmcJ), MmcK (mmck), MmcL (mmcl), MmcM (mmcm), MmcN (mmcn), MmcO (mmco), MmcP (mmcp), MmcQ (mmcq), MmcR (mmcr), MmcS (mmcs), MmcT (mmct), MmcU (mmcu), MmcV (mmcv), Mct (mct), MmcW (mmcw), MmcX (mmcX), and MmcY (mmcy) genes, complete cds; and unknown genes.	Streptomyces lavendulae	66,242	27-MAY-1999

Table 4 (continued)

GB_BA2:AF127374	63734	AF127374	Streptomyces lavendulae LinA homolog, cytochrome P450 hydroxylase ORF4, cytochrome P450 hydroxylase ORF3, MitT (mitT), MitS (mitS), MitR (mitR), MitQ (mitQ), MitP (mitP), MitO (mitO), MitN (mitN), MitM (mitM), MitL (mitL), MitK (mitK), MitJ (mitJ), MitI (mitI), MitH (mitH), MitG (mitG), MitF (mitF), MitE (mitE), MitD (mitD), MitC (mitC), MitB (mitB), MitA (mitA), MmcA (mmcA), MmcB (mmcB), MmcC (mmcC), MmcD (mmcD), MmcE (mmcE), MmcF (mmcF), MmcG (mmcG), MmcH (mmcH), MmcI (mmcI), MmcJ (mmcJ), MmcK (mmcK), MmcL (mmcL), MmcM (mmcM), MmcN (mmcN), MmcO (mmcO), MmcP (mmcP), MmcQ (mmcQ), MmcR (mmcR), MmcS (mmcS), MmcT (mmcT), MmcU (mmcU), MmcV (mmcV), MmcW (mmcW), MmcX (mmcX), and MmcY (mmcY) genes, complete cds; and unknown genes.	Streptomyces lavendulae	38,411	27-MAY-1999
GB_GSS4:AQ741886	742	AQ741886	HS_5569_B2_B02_SP6 RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1145 Col=4 Row=D, genomic survey sequence.	Homo sapiens	38,907	16-Jul-99
GB_EST18:AA567307	741	AA567307	HL01004.5prime HL Drosophila melanogaster head BlueScript Drosophila melanogaster cDNA clone HL01004 5prime, mRNA sequence.	Drosophila melanogaster	38,736	28-Nov-98
GB_EST27:AI402394	630	AI402394	GH21610.5prime GH Drosophila melanogaster head pOT2 Drosophila melanogaster cDNA clone GH21610 5prime, mRNA sequence.	Drosophila melanogaster	41,308	8-Feb-99
GB_GSS10:AQ237646	715	AQ237646	RPCI11-6119.TJB RPCI-11 Homo sapiens genomic clone RPCI-11-6119, genomic survey sequence.	Homo sapiens	44,340	21-Apr-99
GB_EST32:AI726448	562	AI726448	BNLGH5854 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U53418) UDP-glucose dehydrogenase [Glycine max], mRNA sequence.	Gossypium hirsutum	37,003	11-Jun-99
GB_EST32:AI726198	608	AI726198	BNLGH5243 Six-day Cotton fiber Gossypium hirsutum cDNA 5' similar to (U53418) UDP-glucose dehydrogenase [Glycine max], mRNA sequence.	Gossypium hirsutum	40,925	11-Jun-99
GB_PR4:AC002992	154848	AC002992	Homo sapiens chromosome Y, clone 203M13, complete sequence.	Homo sapiens	38,039	13-OCT-1999
GB_EST4:H29653	415	H29653	ym58101.r1 Soares infant brain 1N1B Homo sapiens cDNA clone IMAGE:52678 5' similar to SP:OXDD_BOVIN P31228 D-ASPARTATE OXIDASE ; mRNA sequence.	Homo sapiens	39,036	17-Jul-95
GB_PR3:HSDJ261K5	131974	AL050350	Human DNA sequence from clone 261K5 on chromosome 6q21-22.1. Contains the 3' part of the gene for a novel organic cation transporter (BAC ORF RG331P03), the DDO gene for D-aspartate oxidase (EC 1.4.3.1), ESTs, STSs, GSSs and two putative CpG islands, complete sequence.	Homo sapiens	35,957	23-Nov-99



Table 4 (continued)

GB_EST2:R20147	494	R20147	yg18h02.r1 Soares infant brain 1NIB Homo sapiens cDNA clone IMAGE:32866 5' similar to SP:OXDD_BOVIN P31228 D-ASPARTATE OXIDASE; mRNA sequence.	Homo sapiens	36,437	17-Apr-95
GB_HTG1:CEV6E2	186306	Z96799	Caenorhabditis elegans chromosome V clone Y6E2, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	37,979	02-OCT-1997
GB_HTG1:CEV6E2	186306	Z96799	Caenorhabditis elegans chromosome V clone Y6E2, *** SEQUENCING IN PROGRESS ***; in unordered pieces.	Caenorhabditis elegans	37,979	02-OCT-1997
GB_HTG3:AC011690	72277	AC011690	Homo sapiens clone 17_E_13, LOW-PASS SEQUENCE SAMPLING.	Homo sapiens	35,814	10-OCT-1999
GB_BA1:MSG9883CS	36788	L78828	Mycobacterium leprae cosmid B983 DNA sequence.	Mycobacterium leprae	53,235	15-Jun-96
GB_GSS9:AQ170723	487	AQ170723	HS_2270_B2_F05_MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2270 Col=10 Row=L, genomic survey sequence.	Homo sapiens	39,666	16-OCT-1998
GB_GSS12:AQ349397	791	AQ349397	RPC111-118H16.T.J RPC1-11 Homo sapiens genomic clone RPC1-11- Homo sapiens 118H16, genomic survey sequence.	Homo sapiens	34,204	07-MAY-1999
GB_BA1:MTV026	23740	AL022076	Mycobacterium tuberculosis H37Rv complete genome; segment 157162.	Mycobacterium tuberculosis	37,975	24-Jun-99
GB_IN2:AC005714	177740	AC005714	Drosophila melanogaster, chromosome 2R, region 58D4-58E2, BAC clone BACR48M13, complete sequence.	Drosophila melanogaster	41,226	01-MAY-1999
GB_EST19:AA775050	218	AA775050	ac76e10.s1 Stragene lung (#937210) Homo sapiens cDNA clone IMAGE:868554 3' similar to gb:Y00371_rna1 HEAT SHOCK COGNATE 71 KD PROTEIN (HUMAN); mRNA sequence.	Homo sapiens	40,826	5-Feb-98
GB_BA1:MLCL373	37304	AL035500	Mycobacterium leprae cosmid L373.	Mycobacterium leprae	49,015	27-Aug-99
GB_BA1:MTV044	16150	AL021999	Mycobacterium tuberculosis H37Rv complete genome; segment 45/162.	Mycobacterium tuberculosis	49,192	17-Jun-98
GB_BA1:MLU15180	38675	U15180	Mycobacterium leprae cosmid B1756.	Mycobacterium leprae	45,621	09-MAR-1995
GB_EST37:AI950576	308	AI950576	wx52e08.x1 NCL_CGAP_Lu28 Homo sapiens cDNA clone IMAGE:2547302 3', mRNA sequence.	Homo sapiens	40,909	6-Sep-99
GB_EST37:AI950576	308	AI950576	wx52e08.x1 NCL_CGAP_Lu28 Homo sapiens cDNA clone IMAGE:2547302 3', mRNA sequence.	Homo sapiens	40,288	6-Sep-99
GB_EST34:AV149547	302	AV149547	AV149547 Mus musculus C57BL/6J 10-11 day embryo Mus musculus cDNA clone 2810489D03, mRNA sequence.	Mus musculus	38,627	5-Jul-99
GB_EST35:AV156221	271	AV156221	AV156221 Mus musculus head C57BL/6J 12-day embryo Mus musculus cDNA clone 3000001C24, mRNA sequence.	Mus musculus	33,990	7-Jul-99
GB_EST32:AV054919	274	AV054919	AV054919 Mus musculus pancreas C57BL/6J adult Mus musculus cDNA clone 1810033C08, mRNA sequence.	Mus musculus	36,585	23-Jun-99
GB_BA1:MTV007	32806	AL021184	Mycobacterium tuberculosis H37Rv complete genome; segment 64/162.	Mycobacterium tuberculosis	39,298	17-Jun-98

Table 4 (continued)

GB_BA2:AF027770	30683	AF027770	Mycobacterium smegmatis FxbA (fxbA) gene, partial cds; FxbB (fxbB), FxbC (fxbC), and FxuD (fxtD) genes, complete cds; and unknown genes.	Mycobacterium smegmatis	55,125	03-DEC-1998
GB_BA2:SAU43537	3938	U43537	Streptomyces argillaceus mithramycin resistance determinant, ATP-binding protein (mitrA) and membrane protein (mitrB) genes, complete cds.	Streptomyces argillaceus	46,868	5-Sep-96
GB_BA1:CAJ10319	5368	AJ010319	Corynebacterium glutamicum amtP, glnB, glnD genes and partial ftsY and srp genes.	Corynebacterium glutamicum	100,000	14-MAY-1999
GB_BA1:MTCY338	29372	Z74697	Mycobacterium tuberculosis H37Rv complete genome; segment 127/162.	Mycobacterium tuberculosis	39,785	17-Jun-98
GB_HTG3:AC008733	216140	AC008733	Homo sapiens chromosome 19 clone CITB-E1_2525J15, *** SEQUENCING IN PROGRESS ***; 72 unordered pieces.	Homo sapiens	35,688	3-Aug-99
GB_BA1:BFU64514	3837	U64514	Bacillus firmus dppABC operon, dipeptide transporter protein dppA gene, partial cds, and dipeptide transporter proteins dppB and dppC genes, complete cds.	Bacillus firmus	36,859	1-Feb-97
GB_IN1:CET04C10	20958	Z69885	Caenorhabditis elegans cosmid T04C10, complete sequence.	Caenorhabditis elegans	35,934	2-Sep-99
GB_EST35:AI823090	720	AI823090	L30-944T3 Ice plant Lambda Uni-Zap XR expression library, 30 hours NaCl treatment Mesembryanthemum crystallinum cDNA clone L30-944 5' similar to 60S ribosomal protein L36 (AC004684)[Arabidopsis thaliana], mRNA sequence.	Mesembryanthemum crystallinum	35,770	21-Jul-99
GB_BA1:CJY13333	3315	Y13333	Campylobacter jejuni cfpB gene.	Campylobacter jejuni	53,400	12-Apr-99
GB_BA2:AF065404	181654	AF065404	Bacillus anthracis virulence plasmid PX01, complete sequence.	Bacillus anthracis	45,168	20-OCT-1999
GB_PL2:AC006601	110684	AC006601	Arabidopsis thaliana chromosome V map near 60.5 cM, complete sequence.	Arabidopsis thaliana	36,680	22-Feb-99

### Exemplification

#### **Example 1: Preparation of total genomic DNA of *Corynebacterium glutamicum* ATCC 13032**

- 5           A culture of *Corynebacterium glutamicum* (ATCC 13032) was grown overnight at 30°C with vigorous shaking in BHI medium (Difco). The cells were harvested by centrifugation, the supernatant was discarded and the cells were resuspended in 5 ml buffer-I (5% of the original volume of the culture — all indicated volumes have been calculated for 100 ml of culture volume). Composition of buffer-I: 140.34 g/l sucrose,
- 10   2.46 g/l  $\text{MgSO}_4 \times 7\text{H}_2\text{O}$ , 10 ml/l  $\text{KH}_2\text{PO}_4$  solution (100 g/l, adjusted to pH 6.7 with KOH), 50 ml/l M12 concentrate (10 g/l  $(\text{NH}_4)_2\text{SO}_4$ , 1 g/l NaCl, 2 g/l  $\text{MgSO}_4 \times 7\text{H}_2\text{O}$ , 0.2 g/l  $\text{CaCl}_2$ , 0.5 g/l yeast extract (Difco), 10 ml/l trace-elements-mix (200 mg/l  $\text{FeSO}_4 \times \text{H}_2\text{O}$ , 10 mg/l  $\text{ZnSO}_4 \times 7\text{H}_2\text{O}$ , 3 mg/l  $\text{MnCl}_2 \times 4\text{H}_2\text{O}$ , 30 mg/l  $\text{H}_3\text{BO}_3$ , 20 mg/l  $\text{CoCl}_2 \times 6\text{H}_2\text{O}$ , 1 mg/l  $\text{NiCl}_2 \times 6\text{H}_2\text{O}$ , 3 mg/l  $\text{Na}_2\text{MoO}_4 \times 2\text{H}_2\text{O}$ , 500 mg/l complexing agent
- 15 (EDTA or critic acid), 100 ml/l vitamins-mix (0.2 mg/l biotin, 0.2 mg/l folic acid, 20 mg/l p-amino benzoic acid, 20 mg/l riboflavin, 40 mg/l ca-panthothenate, 140 mg/l nicotinic acid, 40 mg/l pyridoxole hydrochloride, 200 mg/l myo-inositol). Lysozyme was added to the suspension to a final concentration of 2.5 mg/ml. After an approximately 4 h incubation at 37°C, the cell wall was degraded and the resulting
- 20 protoplasts are harvested by centrifugation. The pellet was washed once with 5 ml buffer-I and once with 5 ml TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The pellet was resuspended in 4 ml TE-buffer and 0.5 ml SDS solution (10%) and 0.5 ml NaCl solution (5 M) are added. After adding of proteinase K to a final concentration of 200 µg/ml, the suspension is incubated for ca.18 h at 37°C. The DNA was purified by
- 25 extraction with phenol, phenol-chloroform-isoamylalcohol and chloroform-isoamylalcohol using standard procedures. Then, the DNA was precipitated by adding 1/50 volume of 3 M sodium acetate and 2 volumes of ethanol, followed by a 30 min incubation at -20°C and a 30 min centrifugation at 12,000 rpm in a high speed centrifuge using a SS34 rotor (Sorvall). The DNA was dissolved in 1 ml TE-buffer containing 20
- 30 µg/ml RNaseA and dialysed at 4°C against 1000 ml TE-buffer for at least 3 hours. During this time, the buffer was exchanged 3 times. To aliquots of 0.4 ml of the dialysed DNA solution, 0.4 ml of 2 M LiCl and 0.8 ml of ethanol are added. After a 30

- 111 -

min incubation at -20°C, the DNA was collected by centrifugation (13,000 rpm, Biofuge Fresco, Heraeus, Hanau, Germany). The DNA pellet was dissolved in TE-buffer. DNA prepared by this procedure could be used for all purposes, including southern blotting or construction of genomic libraries.

5

**Example 2: Construction of genomic libraries in *Escherichia coli* of *Corynebacterium glutamicum* ATCC13032.**

Using DNA prepared as described in Example 1, cosmid and plasmid libraries were constructed according to known and well established methods (*see e.g.*, Sambrook, J. *et al.* (1989) "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory Press, or Ausubel, F.M. *et al.* (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.)

Any plasmid or cosmid could be used. Of particular use were the plasmids pBR322 (Sutcliffe, J.G. (1979) *Proc. Natl. Acad. Sci. USA*, 75:3737-3741); pACYC177 (Change & Cohen (1978) *J. Bacteriol* 134:1141-1156), plasmids of the pBS series (pBSSK+, pBSSK- and others; Stratagene, LaJolla, USA), or cosmids as SuperCos1 (Stratagene, LaJolla, USA) or Lorist6 (Gibson, T.J., Rosenthal A. and Waterson, R.H. (1987) *Gene* 53:283-286. Gene libraries specifically for use in *C. glutamicum* may be constructed using plasmid pSL109 (Lee, H.-S. and A. J. Sinskey (1994) *J. Microbiol. Biotechnol.* 4: 256-263).

20

**Example 3: DNA Sequencing and Computational Functional Analysis**

Genomic libraries as described in Example 2 were used for DNA sequencing according to standard methods, in particular by the chain termination method using ABI377 sequencing machines (*see e.g.*, Fleischman, R.D. *et al.* (1995) "Whole-genome Random Sequencing and Assembly of Haemophilus Influenzae Rd., *Science*, 269:496-512). Sequencing primers with the following nucleotide sequences were used: 5'-GGAAACAGTATGACCATG-3' or 5'-GTAAAACGACGGCCAGT-3'.

25

**Example 4: *In vivo* Mutagenesis**

*In vivo* mutagenesis of *Corynebacterium glutamicum* can be performed by passage of plasmid (or other vector) DNA through *E. coli* or other microorganisms (*e.g.* *Bacillus* spp. or yeasts such as *Saccharomyces cerevisiae*) which are impaired in their capabilities to maintain

30

the integrity of their genetic information. Typical mutator strains have mutations in the genes for the DNA repair system (e.g., mutHLS, mutD, mutT, etc.; for reference, see Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, p. 2277-2294, ASM: Washington.) Such strains are well known to those of ordinary skill in the art. The use of such strains is illustrated, for example, in Greener, A. and Callahan, M. (1994) *Strategies* 7: 32-34.

**Example 5: DNA Transfer Between *Escherichia coli* and *Corynebacterium glutamicum***

Several *Corynebacterium* and *Brevibacterium* species contain endogenous plasmids (as e.g., pHM1519 or pBL1) which replicate autonomously (for review see, e.g., Martin, J.F. et al. (1987) *Biotechnology*, 5:137-146). Shuttle vectors for *Escherichia coli* and *Corynebacterium glutamicum* can be readily constructed by using standard vectors for *E. coli* (Sambrook, J. et al. (1989), "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory Press or Ausubel, F.M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons) to which a origin or replication for and a suitable marker from *Corynebacterium glutamicum* is added. Such origins of replication are preferably taken from endogenous plasmids isolated from *Corynebacterium* and *Brevibacterium* species. Of particular use as transformation markers for these species are genes for kanamycin resistance (such as those derived from the Tn5 or Tn903 transposons) or chloramphenicol (Winnacker, E.L. (1987) "From Genes to Clones — Introduction to Gene Technology, VCH, Weinheim). There are numerous examples in the literature of the construction of a wide variety of shuttle vectors which replicate in both *E. coli* and *C. glutamicum*, and which can be used for several purposes, including gene over-expression (for reference, see e.g., Yoshihama, M. et al. (1985) *J. Bacteriol.* 162:591-597, Martin J.F. et al. (1987) *Biotechnology*, 5:137-146 and Eikmanns, B.J. et al. (1991) *Gene*, 102:93-98).

Using standard methods, it is possible to clone a gene of interest into one of the shuttle vectors described above and to introduce such a hybrid vectors into strains of *Corynebacterium glutamicum*. Transformation of *C. glutamicum* can be achieved by protoplast transformation (Kastsumata, R. et al. (1984) *J. Bacteriol.* 159:306-311), electroporation (Liebl, E. et al. (1989) *FEMS Microbiol. Letters*, 53:399-303) and in cases where special vectors are used, also by conjugation (as described e.g. in Schäfer, A et al.

(1990) *J. Bacteriol.* 172:1663-1666). It is also possible to transfer the shuttle vectors for *C. glutamicum* to *E. coli* by preparing plasmid DNA from *C. glutamicum* (using standard methods well-known in the art) and transforming it into *E. coli*. This transformation step can be performed using standard methods, but it is advantageous to use an *Mcr*-deficient  
5 *E. coli* strain, such as NM522 (Gough & Murray (1983) *J. Mol. Biol.* 166:1-19).

Genes may be overexpressed in *C. glutamicum* strains using plasmids which comprise pCG1 (U.S. Patent No. 4,617,267) or fragments thereof, and optionally the gene for kanamycin resistance from TN903 (Grindley, N.D. and Joyce, C.M. (1980) *Proc. Natl. Acad. Sci. USA* 77(12): 7176-7180). In addition, genes may be  
10 overexpressed in *C. glutamicum* strains using plasmid pSL109 (Lee, H.-S. and A. J. Sinskey (1994) *J. Microbiol. Biotechnol.* 4: 256-263).

Aside from the use of replicative plasmids, gene overexpression can also be achieved by integration into the genome. Genomic integration in *C. glutamicum* or other *Corynebacterium* or *Brevibacterium* species may be accomplished by well-known  
15 methods, such as homologous recombination with genomic region(s), restriction endonuclease mediated integration (REMI) (see, *e.g.*, DE Patent 19823834), or through the use of transposons. It is also possible to modulate the activity of a gene of interest by modifying the regulatory regions (*e.g.*, a promoter, a repressor, and/or an enhancer) by sequence modification, insertion, or deletion using site-directed methods (such as  
20 homologous recombination) or methods based on random events (such as transposon mutagenesis or REMI). Nucleic acid sequences which function as transcriptional terminators may also be inserted 3' to the coding region of one or more genes of the invention; such terminators are well-known in the art and are described, for example, in Winnacker, E.L. (1987) *From Genes to Clones – Introduction to Gene Technology*. VCH:  
25 Weinheim.

#### **Example 6: Assessment of the Expression of the Mutant Protein**

Observations of the activity of a mutated protein in a transformed host cell rely on the fact that the mutant protein is expressed in a similar fashion and in a similar quantity to  
30 that of the wild-type protein. A useful method to ascertain the level of transcription of the mutant gene (an indicator of the amount of mRNA available for translation to the gene product) is to perform a Northern blot (for reference see, for example, Ausubel *et al.*

(1988) Current Protocols in Molecular Biology, Wiley: New York), in which a primer designed to bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is extracted, run on gel, transferred to a stable matrix and incubated with this probe, the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information is evidence of the degree of transcription of the mutant gene. Total cellular RNA can be prepared from *Corynebacterium glutamicum* by several methods, all well-known in the art, such as that described in Bormann, E.R. *et al.* (1992) *Mol. Microbiol.* 6: 317-326.

To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a Western blot, may be employed (see, for example, Ausubel *et al.* (1988) Current Protocols in Molecular Biology, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which specifically binds to the desired protein. This probe is generally tagged with a chemiluminescent or colorimetric label which may be readily detected. The presence and quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

#### **Example 7: Growth of Genetically Modified *Corynebacterium glutamicum* — Media and Culture Conditions**

Genetically modified *Corynebacteria* are cultured in synthetic or natural growth media. A number of different growth media for *Corynebacteria* are both well-known and readily available (Lieb *et al.* (1989) *Appl. Microbiol. Biotechnol.*, 32:205-210; von der Osten *et al.* (1998) *Biotechnology Letters*, 11:11-16; Patent DE 4,120,867; Liebl (1992) "The Genus *Corynebacterium*, in: The Prokaryotes, Volume II, Balows, A. *et al.*, eds. Springer-Verlag). These media consist of one or more carbon sources, nitrogen sources, inorganic salts, vitamins and trace elements. Preferred carbon sources are sugars, such as mono-, di-, or polysaccharides. For example, glucose, fructose, mannose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose serve as very good carbon sources. It is also possible to supply sugar to the media via complex compounds such as molasses or other by-products from sugar refinement. It can also be

advantageous to supply mixtures of different carbon sources. Other possible carbon sources are alcohols and organic acids, such as methanol, ethanol, acetic acid or lactic acid. Nitrogen sources are usually organic or inorganic nitrogen compounds, or materials which contain these compounds. Exemplary nitrogen sources include ammonia gas or  
5 ammonia salts, such as  $\text{NH}_4\text{Cl}$  or  $(\text{NH}_4)_2\text{SO}_4$ ,  $\text{NH}_4\text{OH}$ , nitrates, urea, amino acids or complex nitrogen sources like corn steep liquor, soy bean flour, soy bean protein, yeast extract, meat extract and others.

Inorganic salt compounds which may be included in the media include the chloride-, phosphorous- or sulfate- salts of calcium, magnesium, sodium, cobalt,  
10 molybdenum, potassium, manganese, zinc, copper and iron. Chelating compounds can be added to the medium to keep the metal ions in solution. Particularly useful chelating compounds include dihydroxyphenols, like catechol or protocatechuate, or organic acids, such as citric acid. It is typical for the media to also contain other growth factors, such as vitamins or growth promoters, examples of which include biotin, riboflavin, thiamin, folic  
15 acid, nicotinic acid, pantothenate and pyridoxin. Growth factors and salts frequently originate from complex media components such as yeast extract, molasses, corn steep liquor and others. The exact composition of the media compounds depends strongly on the immediate experiment and is individually decided for each specific case. Information about media optimization is available in the textbook "Applied Microbiol. Physiology, A  
20 Practical Approach (*eds.* P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0 19 963577 3). It is also possible to select growth media from commercial suppliers, like standard 1 (Merck) or BHI (grain heart infusion, DIFCO) or others.

All medium components are sterilized, either by heat (20 minutes at 1.5 bar and 121°C) or by sterile filtration. The components can either be sterilized together or, if  
25 necessary, separately. All media components can be present at the beginning of growth, or they can optionally be added continuously or batchwise.

Culture conditions are defined separately for each experiment. The temperature should be in a range between 15°C and 45°C. The temperature can be kept constant or can be altered during the experiment. The pH of the medium should be in the range of 5 to  
30 8.5, preferably around 7.0, and can be maintained by the addition of buffers to the media. An exemplary buffer for this purpose is a potassium phosphate buffer. Synthetic buffers such as MOPS, HEPES, ACES and others can alternatively or simultaneously be used. It



is also possible to maintain a constant culture pH through the addition of NaOH or NH<sub>4</sub>OH during growth. If complex medium components such as yeast extract are utilized, the necessity for additional buffers may be reduced, due to the fact that many complex compounds have high buffer capacities. If a fermentor is utilized for culturing the micro-organisms, the pH can also be controlled using gaseous ammonia.

The incubation time is usually in a range from several hours to several days. This time is selected in order to permit the maximal amount of product to accumulate in the broth. The disclosed growth experiments can be carried out in a variety of vessels, such as microtiter plates, glass tubes, glass flasks or glass or metal fermentors of different sizes.

10 For screening a large number of clones, the microorganisms should be cultured in microtiter plates, glass tubes or shake flasks, either with or without baffles. Preferably 100 ml shake flasks are used, filled with 10% (by volume) of the required growth medium. The flasks should be shaken on a rotary shaker (amplitude 25 mm) using a speed-range of 100 – 300 rpm. Evaporation losses can be diminished by the maintenance

15 of a humid atmosphere; alternatively, a mathematical correction for evaporation losses should be performed.

If genetically modified clones are tested, an unmodified control clone or a control clone containing the basic plasmid without any insert should also be tested. The medium is inoculated to an OD<sub>600</sub> of 0.5 – 1.5 using cells grown on agar plates, such as CM plates

20 (10 g/l glucose, 2,5 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l agar, pH 6.8 with 2M NaOH) that had been incubated at 30°C. Inoculation of the media is accomplished by either introduction of a saline suspension of *C. glutamicum* cells from CM plates or addition of a liquid preculture of this bacterium.

25

### **Example 8 – *In vitro* Analysis of the Function of Mutant Proteins**

The determination of activities and kinetic parameters of enzymes is well established in the art. Experiments to determine the activity of any given altered

30 enzyme must be tailored to the specific activity of the wild-type enzyme, which is well within the ability of one of ordinary skill in the art. Overviews about enzymes in general, as well as specific details concerning structure, kinetics, principles, methods,

- applications and examples for the determination of many enzyme activities may be found, for example, in the following references: Dixon, M., and Webb, E.C., (1979) *Enzymes*. Longmans: London; Fersht, (1985) *Enzyme Structure and Mechanism*. Freeman: New York; Walsh, (1979) *Enzymatic Reaction Mechanisms*. Freeman: San Francisco; Price, N.C., Stevens, L. (1982) *Fundamentals of Enzymology*. Oxford Univ. Press: Oxford; Boyer, P.D., ed. (1983) *The Enzymes*, 3<sup>rd</sup> ed. Academic Press: New York; Bisswanger, H., (1994) *Enzymkinetik*, 2<sup>nd</sup> ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M., eds. (1983-1986) *Methods of Enzymatic Analysis*, 3<sup>rd</sup> ed., vol. I-XII, Verlag Chemie: Weinheim; and Ullmann's
- 10 *Encyclopedia of Industrial Chemistry* (1987) vol. A9, "Enzymes". VCH: Weinheim, p. 352-363.

- The activity of proteins which bind to DNA can be measured by several well-established methods, such as DNA band-shift assays (also called gel retardation assays). The effect of such proteins on the expression of other molecules can be measured using
- 15 reporter gene assays (such as that described in Kolmar, H. *et al.* (1995) *EMBO J.* 14: 3895-3904 and references cited therein). Reporter gene test systems are well known and established for applications in both pro- and eukaryotic cells, using enzymes such as beta-galactosidase, green fluorescent protein, and several others.

- The determination of activity of membrane-transport proteins can be performed
- 20 according to techniques such as those described in Gennis, R.B. (1989) "Pores, Channels and Transporters", in *Biomembranes, Molecular Structure and Function*, Springer: Heidelberg, p. 85-137; 199-234; and 270-322.

#### 25 **Example 9: Analysis of Impact of Mutant Protein on the Production of the Desired Product**

- The effect of the genetic modification in *C. glutamicum* on production of a desired compound (such as an amino acid) can be assessed by growing the modified microorganism under suitable conditions (such as those described above) and analyzing the medium and/or the cellular component for increased production of the desired
- 30 product (*i.e.*, an amino acid). Such analysis techniques are well known to one of ordinary skill in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical

- chromatography such as high performance liquid chromatography (see, for example, Ullman, Encyclopedia of Industrial Chemistry, vol. A2, p. 89-90 and p. 443-613, VCH: Weinheim (1985); Fallon, A. *et al.*, (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm *et al.* (1993) Biotechnology, vol. 3, Chapter III: "Product recovery and purification", page 469-714, VCH: Weinheim; Belter, P.A. *et al.* (1988) Bioseparations: downstream processing for biotechnology, John Wiley and Sons; Kennedy, J.F. and Cabral, J.M.S. (1992) Recovery processes for biological materials, John Wiley and Sons; Shaeiwitz, J.A. and Henry, J.D. (1988) Biochemical separations, in: Ulmann's Encyclopedia of Industrial Chemistry, vol. B3, Chapter 11, page 1-27, VCH: Weinheim; and Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications.)

In addition to the measurement of the final product of fermentation, it is also possible to analyze other components of the metabolic pathways utilized for the production of the desired compound, such as intermediates and side-products, to determine the overall efficiency of production of the compound. Analysis methods include measurements of nutrient levels in the medium (*e.g.*, sugars, hydrocarbons, nitrogen sources, phosphate, and other ions), measurements of biomass composition and growth, analysis of the production of common metabolites of biosynthetic pathways, and measurement of gasses produced during fermentation. Standard methods for these measurements are outlined in Applied Microbial Physiology, A Practical Approach, P.M. Rhodes and P.F. Stanbury, eds., IRL Press, p. 103-129; 131-163; and 165-192 (ISBN: 0199635773) and references cited therein.

#### **Example 10: Purification of the Desired Product from *C. glutamicum* Culture**

Recovery of the desired product from the *C. glutamicum* cells or supernatant of the above-described culture can be performed by various methods well known in the art. If the desired product is not secreted from the cells, the cells can be harvested from the culture by low-speed centrifugation, the cells can be lysed by standard techniques, such as mechanical force or sonication. The cellular debris is removed by centrifugation, and the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from the *C. glutamicum*

cells, then the cells are removed from the culture by low-speed centrifugation, and the supernate fraction is retained for further purification.

The supernatant fraction from either purification method is subjected to chromatography with a suitable resin, in which the desired molecule is either retained on a chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin while the sample is not. Such chromatography steps may be repeated as necessary, using the same or different chromatography resins. One of ordinary skill in the art would be well-versed in the selection of appropriate chromatography resins and in their most efficacious application for a particular molecule to be purified. The purified product may be concentrated by filtration or ultrafiltration, and stored at a temperature at which the stability of the product is maximized.

There are a wide array of purification methods known to the art and the preceding method of purification is not meant to be limiting. Such purification techniques are described, for example, in Bailey, J.E. & Ollis, D.F. *Biochemical Engineering Fundamentals*, McGraw-Hill: New York (1986).

The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, NIRS, enzymatic assay, or microbiologically. Such analysis methods are reviewed in: Patek *et al.* (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova *et al.* (1996) *Biotekhnologiya* 11: 27-32; and Schmidt *et al.* (1998) *Bioprocess Engineer.* 19: 67-70. Ulmann's Encyclopedia of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A. *et al.* (1987) *Applications of HPLC in Biochemistry in: Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17.

#### **Example 11: Analysis of the Gene Sequences of the Invention**

The comparison of sequences and determination of percent homology between two sequences are art-known techniques, and can be accomplished using a mathematical algorithm, such as the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-68, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA*

90:5873-77. Such an algorithm is incorporated into the NBLAST and XBLAST programs (version 2.0) of Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-10. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to HA nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to HA protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul *et al.*, (1997) *Nucleic Acids Res.* 25(17):3389-3402. When utilizing BLAST and Gapped BLAST programs, one of ordinary skill in the art will know how to optimize the parameters of the program (*e.g.*, XBLAST and NBLAST) for the specific sequence being analyzed.

Another example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Meyers and Miller ((1988) *Comput. Appl. Biosci.* 4: 11-17). Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used. Additional algorithms for sequence analysis are known in the art, and include ADVANCE and ADAM. described in Torelli and Robotti (1994) *Comput. Appl. Biosci.* 10:3-5; and FASTA, described in Pearson and Lipman (1988) *P.N.A.S.* 85:2444-8.

The percent homology between two amino acid sequences can also be accomplished using the GAP program in the GCG software package (available at <http://www.gcg.com>), using either a Blossum 62 matrix or a PAM250 matrix, and a gap weight of 12, 10, 8, 6, or 4 and a length weight of 2, 3, or 4. The percent homology between two nucleic acid sequences can be accomplished using the GAP program in the GCG software package, using standard parameters, such as a gap weight of 50 and a length weight of 3.

A comparative analysis of the gene sequences of the invention with those present in Genbank has been performed using techniques known in the art (see, *e.g.*, Bexevanis and Ouellette, eds. (1998) *Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins.* John Wiley and Sons: New York). The gene sequences of the invention

were compared to genes present in Genbank in a three-step process. In a first step, a BLASTN analysis (e.g., a local alignment analysis) was performed for each of the sequences of the invention against the nucleotide sequences present in Genbank, and the top 500 hits were retained for further analysis. A subsequent FASTA search (e.g., a  
5 combined local and global alignment analysis, in which limited regions of the sequences are aligned) was performed on these 500 hits. Each gene sequence of the invention was subsequently globally aligned to each of the top three FASTA hits, using the GAP program in the GCG software package (using standard parameters). In order to obtain correct results, the length of the sequences extracted from Genbank were adjusted to the  
10 length of the query sequences by methods well-known in the art. The results of this analysis are set forth in Table 4. The resulting data is identical to that which would have been obtained had a GAP (global) analysis alone been performed on each of the genes of the invention in comparison with each of the references in Genbank, but required significantly reduced computational time as compared to such a database-wide GAP  
15 (global) analysis. Sequences of the invention for which no alignments above the cutoff values were obtained are indicated on Table 4 by the absence of alignment information. It will further be understood by one of ordinary skill in the art that the GAP alignment homology percentages set forth in Table 4 under the heading "% homology (GAP)" are listed in the European numerical format, wherein a ',' represents a decimal point. For  
20 example, a value of "40,345" in this column represents "40.345%".

#### **Example 12: Construction and Operation of DNA Microarrays**

The sequences of the invention may additionally be used in the construction and application of DNA microarrays (the design, methodology, and uses of DNA arrays are  
25 well known in the art, and are described, for example, in Schena, M. *et al.* (1995) *Science* 270: 467-470; Wodicka, L. *et al.* (1997) *Nature Biotechnology* 15: 1359-1367; DeSaizieu, A. *et al.* (1998) *Nature Biotechnology* 16: 45-48; and DeRisi, J.L. *et al.* (1997) *Science* 278: 680-686).

DNA microarrays are solid or flexible supports consisting of nitrocellulose,  
30 nylon, glass, silicone, or other materials. Nucleic acid molecules may be attached to the surface in an ordered manner. After appropriate labeling, other nucleic acids or nucleic acid mixtures can be hybridized to the immobilized nucleic acid molecules, and the label

may be used to monitor and measure the individual signal intensities of the hybridized molecules at defined regions. This methodology allows the simultaneous quantification of the relative or absolute amount of all or selected nucleic acids in the applied nucleic acid sample or mixture. DNA microarrays, therefore, permit an analysis of the

5 expression of multiple (as many as 6800 or more) nucleic acids in parallel (see, *e.g.*, Schena, M. (1996) *BioEssays* 18(5): 427-431).

The sequences of the invention may be used to design oligonucleotide primers which are able to amplify defined regions of one or more *C. glutamicum* genes by a nucleic acid amplification reaction such as the polymerase chain reaction. The choice

10 and design of the 5' or 3' oligonucleotide primers or of appropriate linkers allows the covalent attachment of the resulting PCR products to the surface of a support medium described above (and also described, for example, Schena, M. *et al.* (1995) *Science* 270: 467-470).

Nucleic acid microarrays may also be constructed by *in situ* oligonucleotide

15 synthesis as described by Wodicka, L. *et al.* (1997) *Nature Biotechnology* 15: 1359-1367. By photolithographic methods, precisely defined regions of the matrix are exposed to light. Protective groups which are photolabile are thereby activated and undergo nucleotide addition, whereas regions that are masked from light do not undergo any modification. Subsequent cycles of protection and light activation permit the

20 synthesis of different oligonucleotides at defined positions. Small, defined regions of the genes of the invention may be synthesized on microarrays by solid phase oligonucleotide synthesis.

The nucleic acid molecules of the invention present in a sample or mixture of nucleotides may be hybridized to the microarrays. These nucleic acid molecules can be

25 labeled according to standard methods. In brief, nucleic acid molecules (*e.g.*, mRNA molecules or DNA molecules) are labeled by the incorporation of isotopically or fluorescently labeled nucleotides, *e.g.*, during reverse transcription or DNA synthesis. Hybridization of labeled nucleic acids to microarrays is described (*e.g.*, in Schena, M. *et al.* (1995) *supra*; Wodicka, L. *et al.* (1997), *supra*; and DeSaizieu A. *et al.* (1998),

30 *supra*). The detection and quantification of the hybridized molecule are tailored to the specific incorporated label. Radioactive labels can be detected, for example, as

described in Schena, M. *et al.* (1995) *supra*) and fluorescent labels may be detected, for example, by the method of Shalon *et al.* (1996) *Genome Research* 6: 639-645).

The application of the sequences of the invention to DNA microarray technology, as described above, permits comparative analyses of different strains of *C. glutamicum* or other Corynebacteria. For example, studies of inter-strain variations based on individual transcript profiles and the identification of genes that are important for specific and/or desired strain properties such as pathogenicity, productivity and stress tolerance are facilitated by nucleic acid array methodologies. Also, comparisons of the profile of expression of genes of the invention during the course of a fermentation reaction are possible using nucleic acid array technology.

### **Example 13: Analysis of the Dynamics of Cellular Protein Populations (Proteomics)**

The genes, compositions, and methods of the invention may be applied to study the interactions and dynamics of populations of proteins, termed 'proteomics'. Protein populations of interest include, but are not limited to, the total protein population of *C. glutamicum* (e.g., in comparison with the protein populations of other organisms), those proteins which are active under specific environmental or metabolic conditions (e.g., during fermentation, at high or low temperature, or at high or low pH), or those proteins which are active during specific phases of growth and development.

Protein populations can be analyzed by various well-known techniques, such as gel electrophoresis. Cellular proteins may be obtained, for example, by lysis or extraction, and may be separated from one another using a variety of electrophoretic techniques. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) separates proteins largely on the basis of their molecular weight. Isoelectric focusing polyacrylamide gel electrophoresis (IEF-PAGE) separates proteins by their isoelectric point (which reflects not only the amino acid sequence but also posttranslational modifications of the protein). Another, more preferred method of protein analysis is the consecutive combination of both IEF-PAGE and SDS-PAGE, known as 2-D-gel electrophoresis (described, for example, in Hermann *et al.* (1998) *Electrophoresis* 19: 3217-3221; Fountoulakis *et al.* (1998) *Electrophoresis* 19: 1193-1202; Langen *et al.* (1997) *Electrophoresis* 18: 1184-1192; Antelmann *et al.* (1997) *Electrophoresis* 18:



1451-1463). Other separation techniques may also be utilized for protein separation, such as capillary gel electrophoresis; such techniques are well known in the art.

Proteins separated by these methodologies can be visualized by standard techniques, such as by staining or labeling. Suitable stains are known in the art, and  
5 include Coomassie Brilliant Blue, silver stain, or fluorescent dyes such as Sypro Ruby (Molecular Probes). The inclusion of radioactively labeled amino acids or other protein precursors (*e.g.*,  $^{35}\text{S}$ -methionine,  $^{35}\text{S}$ -cysteine,  $^{14}\text{C}$ -labelled amino acids,  $^{15}\text{N}$ -amino acids,  $^{15}\text{NO}_3$  or  $^{15}\text{NH}_4^+$  or  $^{13}\text{C}$ -labelled amino acids) in the medium of *C. glutamicum* permits the labeling of proteins from these cells prior to their separation. Similarly,  
10 fluorescent labels may be employed. These labeled proteins can be extracted, isolated and separated according to the previously described techniques.

Proteins visualized by these techniques can be further analyzed by measuring the amount of dye or label used. The amount of a given protein can be determined quantitatively using, for example, optical methods and can be compared to the amount  
15 of other proteins in the same gel or in other gels. Comparisons of proteins on gels can be made, for example, by optical comparison, by spectroscopy, by image scanning and analysis of gels, or through the use of photographic films and screens. Such techniques are well-known in the art.

To determine the identity of any given protein, direct sequencing or other  
20 standard techniques may be employed. For example, N- and/or C-terminal amino acid sequencing (such as Edman degradation) may be used, as may mass spectrometry (in particular MALDI or ESI techniques (see, *e.g.*, Langen *et al.* (1997) *Electrophoresis* 18: 1184-1192)). The protein sequences provided herein can be used for the identification of *C. glutamicum* proteins by these techniques.

25 The information obtained by these methods can be used to compare patterns of protein presence, activity, or modification between different samples from various biological conditions (*e.g.*, different organisms, time points of fermentation, media conditions, or different biotopes, among others). Data obtained from such experiments alone, or in combination with other techniques, can be used for various applications,  
30 such as to compare the behavior of various organisms in a given (*e.g.*, metabolic) situation, to increase the productivity of strains which produce fine chemicals or to increase the efficiency of the production of fine chemicals.

- 125 -

**Equivalents**

Those of ordinary skill in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the

5 following claims.

What is claimed:

1. An isolated nucleic acid molecule from *Corynebacterium glutamicum* encoding an HA protein, or a portion thereof, provided that the nucleic acid molecule does not consist of any of the F-designated genes set forth in Table 1.
2. The isolated nucleic acid molecule of claim 1, wherein said nucleic acid molecule encodes an HA protein involved in the production of a fine chemical.
3. An isolated *Corynebacterium glutamicum* nucleic acid molecule selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, or a portion thereof, provided that the nucleic acid molecule does not consist of any of the F-designated genes set forth in Table 1.
4. An isolated nucleic acid molecule which encodes a polypeptide sequence selected from the group consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, provided that the nucleic acid molecule does not consist of any of the F-designated genes set forth in Table 1.
5. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide selected from the group of amino acid sequences consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, provided that the nucleic acid molecule does not consist of any of the F-designated genes set forth in Table 1.
6. An isolated nucleic acid molecule comprising a nucleotide sequence which is at least 50% homologous to a nucleotide sequence selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, or a portion thereof, provided that the nucleic acid molecule does not consist of any of the F-designated genes set forth in Table 1.

- 127 -

7. An isolated nucleic acid molecule comprising a fragment of at least 15 nucleotides of a nucleic acid comprising a nucleotide sequence selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, provided that the nucleic acid molecule does not consist of any of the F-designated genes set forth in Table 1.
8. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule of any one of claims 1-7 under stringent conditions.
9. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1-8 or a portion thereof and a nucleotide sequence encoding a heterologous polypeptide.
10. A vector comprising the nucleic acid molecule of any one of claims 1-9.
11. The vector of claim 10, which is an expression vector.
12. A host cell transfected with the expression vector of claim 11.
13. The host cell of claim 12, wherein said cell is a microorganism.
14. The host cell of claim 13, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.
15. The host cell of claim 12, wherein the expression of said nucleic acid molecule results in the modulation in production of a fine chemical from said cell.
16. The host cell of claim 15, wherein said fine chemical is selected from the group consisting of: organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides, and enzymes.

17. A method of producing a polypeptide comprising culturing the host cell of claim 12 in an appropriate culture medium to, thereby, produce the polypeptide.
- 5 18. An isolated HA polypeptide from *Corynebacterium glutamicum*, or a portion thereof.
19. The polypeptide of claim 18, wherein said polypeptide is involved in the production of a fine chemical production.
- 10 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, provided that the amino acid sequence is not encoded by any of the F-designated genes set forth in Table 1.
- 15 21. An isolated polypeptide comprising a naturally occurring allelic variant of a polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth in as even-numbered SEQ ID NOs of the Sequence Listing, or a portion thereof, provided that the amino acid sequence is not encoded
- 20 by any of the F-designated genes set forth in Table 1.
22. The isolated polypeptide of any of claims 18-21, further comprising heterologous amino acid sequences.
- 25 23. An isolated polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 50% homologous to a nucleic acid selected from the group consisting of those sequences set forth as odd-numbered SEQ ID NOs of the Sequence Listing, provided that the nucleic acid molecule does not consist of any of the F-designated nucleic acid molecules set forth in Table 1.
- 30 24. An isolated polypeptide comprising an amino acid sequence which is at least 50% homologous to an amino acid sequence selected from the group consisting of those

- 129 -

sequences set forth as even-numbered SEQ ID NOs of the Sequence Listing, provided that the amino acid sequence is not encoded by any of the F-designated genes set forth in Table 1.

- 5 25. A method for producing a fine chemical, comprising culturing a cell containing a vector of claim 12 such that the fine chemical is produced.
26. The method of claim 25, wherein said method further comprises the step of recovering the fine chemical from said culture.
- 10 27. The method of claim 25, wherein said method further comprises the step of transfecting said cell with the vector of claim 11 to result in a cell containing said vector.
- 15 28. The method of claim 25, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.
29. The method of claim 25, wherein said cell is selected from the group consisting of: *Corynebacterium glutamicum*, *Corynebacterium herculis*, *Corynebacterium lilium*,  
20 *Corynebacterium acetoacidophilum*, *Corynebacterium acetoglutamicum*,  
*Corynebacterium acetophilum*, *Corynebacterium ammoniagenes*, *Corynebacterium fujiokense*, *Corynebacterium nitrilophilus*, *Brevibacterium ammoniagenes*,  
*Brevibacterium butanicum*, *Brevibacterium divaricatum*, *Brevibacterium flavum*,  
*Brevibacterium healii*, *Brevibacterium ketoglutamicum*, *Brevibacterium*  
25 *ketosoreductum*, *Brevibacterium lactofermentum*, *Brevibacterium linens*,  
*Brevibacterium paraffinolyticum*, and those strains set forth in Table 3.
30. The method of claim 25, wherein expression of the nucleic acid molecule from said vector results in modulation of production of said fine chemical.
- 30 31. The method of claim 25, wherein said fine chemical is selected from the group consisting of: organic acids, proteinogenic and nonproteinogenic amino acids, purine

- 130 -

and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, polyketides, and enzymes.

5 32. The method of claim 25, wherein said fine chemical is an amino acid.

33. The method of claim 32, wherein said amino acid is drawn from the group consisting of: lysine, glutamate, glutamine, alanine, aspartate, glycine, serine, threonine, methionine, cysteine, valine, leucine, isoleucine, arginine, proline, histidine,  
10 tyrosine, phenylalanine, and tryptophan.

34. A method for producing a fine chemical, comprising culturing a cell whose genomic DNA has been altered by the inclusion of a nucleic acid molecule of any one of claims 1-9.

15

35. A method for diagnosing the presence or activity of *Corynebacterium diphtheriae* in a subject, comprising detecting the presence of one or more of SEQ ID NOs 1 through 440 of the Sequence Listing in the subject, provided that the sequences are not or are not encoded by any of the F-designated sequences set forth in Table 1, thereby diagnosing  
20 the presence or activity of *Corynebacterium diphtheriae* in the subject.

36. A host cell comprising a nucleic acid molecule selected from the group consisting of the nucleic acid molecules set forth as odd-numbered SEQ ID NOs of the Sequence Listing, wherein the nucleic acid molecule is disrupted.

25

37. A host cell comprising a nucleic acid molecule selected from the group consisting of the nucleic acid molecules set forth as odd-numbered SEQ ID NOs in the Sequence Listing, wherein the nucleic acid molecule comprises one or more nucleic acid modifications from the sequence set forth as odd-numbered SEQ ID NOs of the  
30 Sequence Listing s.

- 131 -

38. A host cell comprising a nucleic acid molecule selected from the group consisting of the nucleic acid molecules set forth as odd-numbered SEQ ID NOs of the Sequence Listing, wherein the regulatory region of the nucleic acid molecule is modified relative to the wild-type regulatory region of the molecule.



## SEQUENCE LISTING

<110> BASF Aktiengesellschaft  
 <120> CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
 INVOLVED IN HOMEOSTASIS AND ADAPTATION  
 <130> BGI-128CPPC  
 <140>  
 <141>  
 <160> 440  
  
 <210> 1  
 <211> 314  
 <212> DNA  
 <213> Corynebacterium glutamicum  
  
 <220>  
 <221> CDS  
 <222> (1)..(291)  
 <223> RXA02548  
  
 <400> 1  
 cca ccg atc tac ttc tcc cac gac cgc gaa gtt ttc gag cgc gac ggc 48  
 Pro Pro Ile Tyr Phe Ser His Asp Arg Glu Val Phe Glu Arg Asp Gly  
 1 5 10 15  
  
 atg tgg ctg acc gca ggc gag tgg ggt gga cca aag aag ggc gag gag 96  
 Met Trp Leu Thr Ala Gly Glu Trp Gly Gly Pro Lys Lys Gly Glu Glu  
 20 25 30  
  
 atc gtc acc aag act gtc cgc tac cgc acc gtc ggc gat atg tcc tgc 144  
 Ile Val Thr Lys Thr Val Arg Tyr Arg Thr Val Gly Asp Met Ser Cys  
 35 40 45  
  
 acc ggt gct gtg ctc tcc gaa gcc cgc acc att gac gat gtg atc gaa 192  
 Thr Gly Ala Val Leu Ser Glu Ala Arg Thr Ile Asp Asp Val Ile Glu  
 50 55 60  
  
 gag atc gcc acc tcc acc ctt acc gaa cgt ggc gca acc cgc gcc gat 240  
 Glu Ile Ala Thr Ser Thr Leu Thr Glu Arg Gly Ala Thr Arg Ala Asp  
 65 70 75 80  
  
 gac cgc ctc agc gaa tcc gca atg gaa gac cgc aag aag gaa ggc tac 288  
 Asp Arg Leu Ser Glu Ser Ala Met Glu Asp Arg Lys Lys Glu Gly Tyr  
 85 90 95  
  
 ttc tgatgactgc tccaaccttg aat 314  
 Phe  
  
 <210> 2  
 <211> 97  
 <212> PRT  
 <213> Corynebacterium glutamicum  
  
 <400> 2  
 Pro Pro Ile Tyr Phe Ser His Asp Arg Glu Val Phe Glu Arg Asp Gly  
 1 5 10 15  
  
 Met Trp Leu Thr Ala Gly Glu Trp Gly Gly Pro Lys Lys Gly Glu Glu  
 20 25 30

Ile Val Thr Lys Thr Val Arg Tyr Arg Thr Val Gly Asp Met Ser Cys  
                   35                                  40                                  45  
 Thr Gly Ala Val Leu Ser Glu Ala Arg Thr Ile Asp Asp Val Ile Glu  
                   50                                  55                                  60  
 Glu Ile Ala Thr Ser Thr Leu Thr Glu Arg Gly Ala Thr Arg Ala Asp  
                   65                                  70                                  75                                  80  
 Asp Arg Leu Ser Glu Ser Ala Met Glu Asp Arg Lys Lys Glu Gly Tyr  
                                   85                                  90                                  95

Phe

<210> 3  
 <211> 980  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (1)..(957)  
 <223> RXN00249

<400> 3  
 acc ggc gtc tcc acc tcc cag gtt gta gtt ttg ctt gtc gac gcc cgc 48  
 Thr Gly Val Ser Thr Ser Gln Val Val Val Leu Leu Val Asp Ala Arg  
                   1                                  5                                  10                                  15  
 cac ggc gtc gtc gag cag acc cgc cgc cac ctg tcc gta tcg gct ctg 96  
 His Gly Val Val Glu Gln Thr Arg Arg His Leu Ser Val Ser Ala Leu  
                                   20                                  25                                  30  
 ctg ggc gta cgc acg gtg atc ctc gca gtc aac aaa att gac ctt gtt 144  
 Leu Gly Val Arg Thr Val Ile Leu Ala Val Asn Lys Ile Asp Leu Val  
                   35                                  40                                  45  
 gat tac agc gaa gaa gtc ttc cgc aac att gaa aag gaa ttc gtt ggc 192  
 Asp Tyr Ser Glu Glu Val Phe Arg Asn Ile Glu Lys Glu Phe Val Gly  
                   50                                  55                                  60  
 ctg gca tct gca ctt gat gtc aca gac acc cac gtt gtt cca atc tct 240  
 Leu Ala Ser Ala Leu Asp Val Thr Asp Thr His Val Val Pro Ile Ser  
                   65                                  70                                  75                                  80  
 gcg ctc aag ggc gac aac gtt gca gaa cct tcc acc cac atg gat tgg 288  
 Ala Leu Lys Gly Asp Asn Val Ala Glu Pro Ser Thr His Met Asp Trp  
                                   85                                  90                                  95  
 tac acc gga cca acc gtg ctg gaa atc ctg gaa aac gta gaa gtt tcc 336  
 Tyr Thr Gly Pro Thr Val Leu Glu Ile Leu Glu Asn Val Glu Val Ser  
                   100                                  105                                  110  
 cac ggc cgt gca cac gac ctg ggc ttc cgc ttc cca atc cag tac gtc 384  
 His Gly Arg Ala His Asp Leu Gly Phe Arg Phe Pro Ile Gln Tyr Val  
                   115                                  120                                  125  
 atc cgc gag cac gcc acc gac tac cgt ggc tac gcc ggc acc atc aac 432

```

Ile Arg Glu His Ala Thr Asp Tyr Arg Gly Tyr Ala Gly Thr Ile Asn
130                               135                               140

gct ggt tcc gtc tcc gtg ggc gat acc gtg tac cta cct gaa ggc cgc 480
Ala Gly Ser Val Ser Val Gly Asp Thr Val Tyr Leu Pro Glu Gly Arg
145                               150                               155                               160

acc acc cag gtc acc cac atc gat tcc gct gac gga tcc ctc cag acc 528
Thr Thr Gln Val Thr His Ile Asp Ser Ala Asp Gly Ser Leu Gln Thr
165                               170                               175

gca tca gtt gga gaa gcc gtt gtc ctg cgc cta gcc cag gaa atc gac 576
Ala Ser Val Gly Glu Ala Val Val Leu Arg Leu Ala Gln Glu Ile Asp
180                               185                               190

ctc atc cgc ggc gaa ctc atc gct ggc gaa gac cgc cca gaa tcc gtt 624
Leu Ile Arg Gly Glu Leu Ile Ala Gly Glu Asp Arg Pro Glu Ser Val
195                               200                               205

cgc tcc ttc aac gcc act gtt gtt ggc ttg gcc gat cgc acc atc aaa 672
Arg Ser Phe Asn Ala Thr Val Val Gly Leu Ala Asp Arg Thr Ile Lys
210                               215                               220

cca ggt gca gca gtc aag gtt cgc tac ggc acc gag ctg gtc cgc gga 720
Pro Gly Ala Ala Val Lys Val Arg Tyr Gly Thr Glu Leu Val Arg Gly
225                               230                               235

cgc gtc gca gcc atc gaa cga gtc ctc gac atc gac ggc gtc aac gac 768
Arg Val Ala Ala Ile Glu Arg Val Leu Asp Ile Asp Gly Val Asn Asp
245                               250                               255

aac gaa gca cca gaa acc tac ggc ctc aac gac atc gca cac gtg cgc 816
Asn Glu Ala Pro Glu Thr Tyr Gly Leu Asn Asp Ile Ala His Val Arg
260                               265                               270

atc gac gtt gca ggc gaa ctc gaa gtt gaa gat tac gct gcc cgc ggc 864
Ile Asp Val Ala Gly Glu Leu Glu Val Glu Asp Tyr Ala Ala Arg Gly
275                               280                               285

gcc atc gga tcc ttc ctc ctc atc gac caa tcc tcc ggc gat acc ctc 912
Ala Ile Gly Ser Phe Leu Leu Ile Asp Gln Ser Ser Gly Asp Thr Leu
290                               295                               300

gca gct ggc ttg gtt ggc cac cgc cta cgc aat aac tgg tcg atc 957
Ala Ala Gly Leu Val Gly His Arg Leu Arg Asn Asn Trp Ser Ile
305                               310                               315

tagaccagtg tcttaggcaa gac 980

```

<210> 4  
 <211> 319  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 4  
 Thr Gly Val Ser Thr Ser Gln Val Val Val Leu Leu Val Asp Ala Arg  
 1 5 10 15  
 His Gly Val Val Glu Gln Thr Arg Arg His Leu Ser Val Ser Ala Leu  
 20 25 30

Leu Gly Val Arg Thr Val Ile Leu Ala Val Asn Lys Ile Asp Leu Val  
 35 40 45  
 Asp Tyr Ser Glu Glu Val Phe Arg Asn Ile Glu Lys Glu Phe Val Gly  
 50 55 60  
 Leu Ala Ser Ala Leu Asp Val Thr Asp Thr His Val Val Pro Ile Ser  
 65 70 75 80  
 Ala Leu Lys Gly Asp Asn Val Ala Glu Pro Ser Thr His Met Asp Trp  
 85 90 95  
 Tyr Thr Gly Pro Thr Val Leu Glu Ile Leu Glu Asn Val Glu Val Ser  
 100 105 110  
 His Gly Arg Ala His Asp Leu Gly Phe Arg Phe Pro Ile Gln Tyr Val  
 115 120 125  
 Ile Arg Glu His Ala Thr Asp Tyr Arg Gly Tyr Ala Gly Thr Ile Asn  
 130 135 140  
 Ala Gly Ser Val Ser Val Gly Asp Thr Val Tyr Leu Pro Glu Gly Arg  
 145 150 155 160  
 Thr Thr Gln Val Thr His Ile Asp Ser Ala Asp Gly Ser Leu Gln Thr  
 165 170 175  
 Ala Ser Val Gly Glu Ala Val Val Leu Arg Leu Ala Gln Glu Ile Asp  
 180 185 190  
 Leu Ile Arg Gly Glu Leu Ile Ala Gly Glu Asp Arg Pro Glu Ser Val  
 195 200 205  
 Arg Ser Phe Asn Ala Thr Val Val Gly Leu Ala Asp Arg Thr Ile Lys  
 210 215 220  
 Pro Gly Ala Ala Val Lys Val Arg Tyr Gly Thr Glu Leu Val Arg Gly  
 225 230 235 240  
 Arg Val Ala Ala Ile Glu Arg Val Leu Asp Ile Asp Gly Val Asn Asp  
 245 250 255  
 Asn Glu Ala Pro Glu Thr Tyr Gly Leu Asn Asp Ile Ala His Val Arg  
 260 265 270  
 Ile Asp Val Ala Gly Glu Leu Glu Val Glu Asp Tyr Ala Ala Arg Gly  
 275 280 285  
 Ala Ile Gly Ser Phe Leu Leu Ile Asp Gln Ser Ser Gly Asp Thr Leu  
 290 295 300  
 Ala Ala Gly Leu Val Gly His Arg Leu Arg Asn Asn Trp Ser Ile  
 305 310 315

&lt;210&gt; 5

&lt;211&gt; 977

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(954)

&lt;223&gt; FRXA00249

&lt;400&gt; 5

ggt gtt ttc acc ttc caa ggt gta gtt ttg ctt gtc gac gcc cgc cac	48
Gly Val Phe Thr Phe Gln Gly Val Val Leu Leu Val Asp Ala Arg His	
1 5 10 15	
ggc gtc gtc gag cag acc cgc cgc cac ctg tcc gta tcg gct ctg ctg	96
Gly Val Val Glu Gln Thr Arg Arg His Leu Ser Val Ser Ala Leu Leu	
20 25 30	
ggc gta cgc acg gtg atc ctc gca gtc aac aaa att gac ctt gtt gat	144
Gly Val Arg Thr Val Ile Leu Ala Val Asn Lys Ile Asp Leu Val Asp	
35 40 45	
tac agc gaa gaa gtc ttc cgc aac att gaa aag gaa ttc gtt ggc ctg	192
Tyr Ser Glu Glu Val Phe Arg Asn Ile Glu Lys Glu Phe Val Gly Leu	
50 55 60	
gca tct gca ctt gat gtc aca gac acc cac gtt gtt cca atc tct gcg	240
Ala Ser Ala Leu Asp Val Thr Asp Thr His Val Val Pro Ile Ser Ala	
65 70 75 80	
ctc aag ggc gac aac gtt gca gaa cct tcc acc cac atg gat tgg tac	288
Leu Lys Gly Asp Asn Val Ala Glu Pro Ser Thr His Met Asp Trp Tyr	
85 90 95	
acc gga cca acc gtg ctg gaa atc ctg gaa aac gta gaa gtt tcc cac	336
Thr Gly Pro Thr Val Leu Glu Ile Leu Glu Asn Val Glu Val Ser His	
100 105 110	
ggc cgt gca cac gac ctg ggc ttc cgc ttc cca atc cag tac gtc atc	384
Gly Arg Ala His Asp Leu Gly Phe Arg Phe Pro Ile Gln Tyr Val Ile	
115 120 125	
cgc gag cac gcc acc gac tac cgt ggc tac gcc ggc acc atc aac gct	432
Arg Glu His Ala Thr Asp Tyr Arg Gly Tyr Ala Gly Thr Ile Asn Ala	
130 135 140	
ggt tcc gtc tcc gtg ggc gat acc gtg tac cta cct gaa ggc cgc acc	480
Gly Ser Val Ser Val Gly Asp Thr Val Tyr Leu Pro Glu Gly Arg Thr	
145 150 155 160	
acc cag gtc acc cac atc gat tcc gct gac gga tcc ctc cag acc gca	528
Thr Gln Val Thr His Ile Asp Ser Ala Asp Gly Ser Leu Gln Thr Ala	
165 170 175	
tca gtt gga gaa gcc gtt gtc ctg cgc cta gcc cag gaa atc gac ctc	576
Ser Val Gly Glu Ala Val Val Leu Arg Leu Ala Gln Glu Ile Asp Leu	
180 185 190	
atc cgc ggc gaa ctc atc gct ggc gaa gac cgc cca gaa tcc gtt cgc	624
Ile Arg Gly Glu Leu Ile Ala Gly Glu Asp Arg Pro Glu Ser Val Arg	
195 200 205	
tcc ttc aac gcc act gtt gtt ggc ttg gcc gat cgc acc atc aaa cca	672
Ser Phe Asn Ala Thr Val Val Gly Leu Ala Asp Arg Thr Ile Lys Pro	
210 215 220	

ggt gca gca gtc aag gtt cgc tac ggc acc gag ctg gtc cgc gga cgc 720  
 Gly Ala Ala Val Lys Val Arg Tyr Gly Thr Glu Leu Val Arg Gly Arg  
 225 230 235 240  
 gtc gca gcc atc gaa cga gtc ctc gac atc gac ggc gtc aac gac aac 768  
 Val Ala Ala Ile Glu Arg Val Leu Asp Ile Asp Gly Val Asn Asp Asn  
 245 250 255  
 gaa gca cca gaa acc tac ggc ctc aac gac atc gca cac gtg cgc atc 816  
 Glu Ala Pro Glu Thr Tyr Gly Leu Asn Asp Ile Ala His Val Arg Ile  
 260 265 270  
 gac gtt gca ggc gaa ctc gaa gtt gaa gat tac gct gcc cgc ggc gcc 864  
 Asp Val Ala Gly Glu Leu Glu Val Glu Asp Tyr Ala Ala Arg Gly Ala  
 275 280 285  
 atc gga tcc ttc ctc ctc atc gac caa tcc tcc ggc gat acc ctc gca 912  
 Ile Gly Ser Phe Leu Leu Ile Asp Gln Ser Ser Gly Asp Thr Leu Ala  
 290 295 300  
 gct ggc ttg gtt ggc cac cgc cta cgc aat aac tgg tog atc 954  
 Ala Gly Leu Val Gly His Arg Leu Arg Asn Asn Trp Ser Ile  
 305 310 315  
 tagaccagtg tcttaggcaa gac 977

<210> 6  
 <211> 318  
 <212> PRT  
 <213> *Corynebacterium glutamicum*

<400> 6  
 Gly Val Phe Thr Phe Gln Gly Val Val Leu Leu Val Asp Ala Arg His  
 1 5 10 15  
 Gly Val Val Glu Gln Thr Arg Arg His Leu Ser Val Ser Ala Leu Leu  
 20 25 30  
 Gly Val Arg Thr Val Ile Leu Ala Val Asn Lys Ile Asp Leu Val Asp  
 35 40 45  
 Tyr Ser Glu Glu Val Phe Arg Asn Ile Glu Lys Glu Phe Val Gly Leu  
 50 55 60  
 Ala Ser Ala Leu Asp Val Thr Asp Thr His Val Val Pro Ile Ser Ala  
 65 70 75 80  
 Leu Lys Gly Asp Asn Val Ala Glu Pro Ser Thr His Met Asp Trp Tyr  
 85 90 95  
 Thr Gly Pro Thr Val Leu Glu Ile Leu Glu Asn Val Glu Val Ser His  
 100 105 110  
 Gly Arg Ala His Asp Leu Gly Phe Arg Phe Pro Ile Gln Tyr Val Ile  
 115 120 125  
 Arg Glu His Ala Thr Asp Tyr Arg Gly Tyr Ala Gly Thr Ile Asn Ala  
 130 135 140

Gly Ser Val Ser Val Gly Asp Thr Val Tyr Leu Pro Glu Gly Arg Thr  
 145 150 155 160  
 Thr Gln Val Thr His Ile Asp Ser Ala Asp Gly Ser Leu Gln Thr Ala  
 165 170 175  
 Ser Val Gly Glu Ala Val Val Leu Arg Leu Ala Gln Glu Ile Asp Leu  
 180 185 190  
 Ile Arg Gly Glu Leu Ile Ala Gly Glu Asp Arg Pro Glu Ser Val Arg  
 195 200 205  
 Ser Phe Asn Ala Thr Val Val Gly Leu Ala Asp Arg Thr Ile Lys Pro  
 210 215 220  
 Gly Ala Ala Val Lys Val Arg Tyr Gly Thr Glu Leu Val Arg Gly Arg  
 225 230 235 240  
 Val Ala Ala Ile Glu Arg Val Leu Asp Ile Asp Gly Val Asn Asp Asn  
 245 250 255  
 Glu Ala Pro Glu Thr Tyr Gly Leu Asn Asp Ile Ala His Val Arg Ile  
 260 265 270  
 Asp Val Ala Gly Glu Leu Glu Val Glu Asp Tyr Ala Ala Arg Gly Ala  
 275 280 285  
 Ile Gly Ser Phe Leu Leu Ile Asp Gln Ser Ser Gly Asp Thr Leu Ala  
 290 295 300  
 Ala Gly Leu Val Gly His Arg Leu Arg Asn Asn Trp Ser Ile  
 305 310 315

<210> 7  
 <211> 954  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(931)  
 <223> RXA01073

<400> 7  
 taaccgactc cagcactaaa ctccaaaccc ttggcccgca ccgccaaagt ttagcgcgcc 60  
 ccaagacacc accgcgccat gtttgcctag gattaggtac atg aca aac act caa 115  
 Met Thr Asn Thr Gln  
 1 5  
 acc gag atc att aat gaa cta aag gtg agc cca gca atc gac gtg gcc 163  
 Thr Glu Ile Ile Asn Glu Leu Lys Val Ser Pro Ala Ile Asp Val Ala  
 10 15 20  
 aag gaa gtt gaa ttc cgt gtg cag ttc ctc gtc gat tac ctg cgg gct 211  
 Lys Glu Val Glu Phe Arg Val Gln Phe Leu Val Asp Tyr Leu Arg Ala  
 25 30 35  
 tcc cat aca aaa ggc ttt gtt ctt ggt att tca ggt ggc cag gat tcc 259  
 Ser His Thr Lys Gly Phe Val Leu Gly Ile Ser Gly Gly Gln Asp Ser

40	45	50	
act ctt gcg gga cga ctc acg cag ctg gca gta gag cgc att cgt gcg Thr Leu Ala Gly Arg Leu Thr Gln Leu Ala Val Glu Arg Ile Arg Ala 55 60 65			307
gaa gaa aac agc acg gat tat gtc ttc tac gca gtt cgc ctc ccc tac Glu Glu Asn Ser Thr Asp Tyr Val Phe Tyr Ala Val Arg Leu Pro Tyr 70 75 80 85			355
gcg atc cag gca gat gag gac gat gcg caa gtt gca ttg gaa ttc atc Ala Ile Gln Ala Asp Glu Asp Asp Ala Gln Val Ala Leu Glu Phe Ile 90 95 100			403
gca cct gac aag agc gtg acc gtc aac gtt aaa gac gca acg gac gcc Ala Pro Asp Lys Ser Val Thr Val Asn Val Lys Asp Ala Thr Asp Ala 105 110 115			451
acc gaa gca act gtt gca gct gct ttg gaa ctt cct gag ctg acc gac Thr Glu Ala Thr Val Ala Ala Ala Leu Glu Leu Pro Glu Leu Thr Asp 120 125 130			499
ttc aat cgg ggc aat att aaa gct cgc caa cgc atg gtt gcc cag tac Phe Asn Arg Gly Asn Ile Lys Ala Arg Gln Arg Met Val Ala Gln Tyr 135 140 145			547
gca atc gca ggc cag ttg ggc ttg ctg gtt att ggc act gat cac gcg Ala Ile Ala Gly Gln Leu Gly Leu Leu Val Ile Gly Thr Asp His Ala 150 155 160 165			595
gct gaa aac gtc acg ggg ttc ttc acc aaa ttc ggt gat ggc gca gct Ala Glu Asn Val Thr Gly Phe Phe Thr Lys Phe Gly Asp Gly Ala Ala 170 175 180			643
gac ctg ctt cct ttg gca ggt ttg agc aag cgt caa gga gct gcc att Asp Leu Leu Pro Leu Ala Gly Leu Ser Lys Arg Gln Gly Ala Ala Ile 185 190 195			691
ctg gag cac ctg ggt gca cct tca agc acg tgg acc aag gtt cct acc Leu Glu His Leu Gly Ala Pro Ser Ser Thr Trp Thr Lys Val Pro Thr 200 205 210			739
gct gat ttg gaa gag gat cgc cca gcg ttg cca gat gag gaa gca ctt Ala Asp Leu Glu Glu Asp Arg Pro Ala Leu Pro Asp Glu Glu Ala Leu 215 220 225			787
ggt gtg tcg tat gcg gac atc gat aat tac ctg gaa aac aag ccc gat Gly Val Ser Tyr Ala Asp Ile Asp Asn Tyr Leu Glu Asn Lys Pro Asp 230 235 240 245			835
gtc agt gaa aaa gcc cag cag cgc att gag cac ctg tgg aag gtg ggc Val Ser Glu Lys Ala Gln Gln Arg Ile Glu His Leu Trp Lys Val Gly 250 255 260			883
cag cac aag cgc cac ctc cct gct acc ccg cag gaa aat tgg tgg cgt Gln His Lys Arg His Leu Pro Ala Thr Pro Gln Glu Asn Trp Trp Arg 265 270 275			931
taatccaaca gtttgagtgt cgc			954



&lt;210&gt; 8

&lt;211&gt; 277

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 8

```

Met Thr Asn Thr Gln Thr Glu Ile Ile Asn Glu Leu Lys Val Ser Pro
  1           5           10           15

Ala Ile Asp Val Ala Lys Glu Val Glu Phe Arg Val Gln Phe Leu Val
      20           25           30

Asp Tyr Leu Arg Ala Ser His Thr Lys Gly Phe Val Leu Gly Ile Ser
      35           40           45

Gly Gly Gln Asp Ser Thr Leu Ala Gly Arg Leu Thr Gln Leu Ala Val
      50           55           60

Glu Arg Ile Arg Ala Glu Glu Asn Ser Thr Asp Tyr Val Phe Tyr Ala
      65           70           75           80

Val Arg Leu Pro Tyr Ala Ile Gln Ala Asp Glu Asp Asp Ala Gln Val
      85           90           95

Ala Leu Glu Phe Ile Ala Pro Asp Lys Ser Val Thr Val Asn Val Lys
      100          105          110

Asp Ala Thr Asp Ala Thr Glu Ala Thr Val Ala Ala Ala Leu Glu Leu
      115          120          125

Pro Glu Leu Thr Asp Phe Asn Arg Gly Asn Ile Lys Ala Arg Gln Arg
      130          135          140

Met Val Ala Gln Tyr Ala Ile Ala Gly Gln Leu Gly Leu Leu Val Ile
      145          150          155          160

Gly Thr Asp His Ala Ala Glu Asn Val Thr Gly Phe Phe Thr Lys Phe
      165          170          175

Gly Asp Gly Ala Ala Asp Leu Leu Pro Leu Ala Gly Leu Ser Lys Arg
      180          185          190

Gln Gly Ala Ala Ile Leu Glu His Leu Gly Ala Pro Ser Ser Thr Trp
      195          200          205

Thr Lys Val Pro Thr Ala Asp Leu Glu Glu Asp Arg Pro Ala Leu Pro
      210          215          220

Asp Glu Glu Ala Leu Gly Val Ser Tyr Ala Asp Ile Asp Asn Tyr Leu
      225          230          235          240

Glu Asn Lys Pro Asp Val Ser Glu Lys Ala Gln Gln Arg Ile Glu His
      245          250          255

Leu Trp Lys Val Gly Gln His Lys Arg His Leu Pro Ala Thr Pro Gln
      260          265          270

Glu Asn Trp Trp Arg
      275

```

```
<220>  
<221> CDS  
<222> (101)..(586)  
<223> RXN02913
```

<210> 10

<211> 162  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 10

```

Met Ile Pro Gly Glu Tyr Ile Leu Ser Ser Glu Ser Leu Thr Gly Asn
 1           5           10           15
Val Gly Arg Glu Ala Lys Thr Ile Glu Ile Ile Asn Thr Gly Asp Arg
          20           25           30
Pro Val Gln Ile Gly Ser His Phe His Phe Ala Glu Val Asn Pro Ser
          35           40           45
Ile Ser Phe Asp Arg Ser Glu Gly Tyr Gly Phe Arg Leu Asp Ile Pro
          50           55           60
Ser Gly Thr Ala Val Arg Leu Glu Pro Gly Asp Ala Arg Thr Val Asn
          65           70           75           80
Leu Val Ala Ile Gly Gly Asp Arg Ile Val Ala Gly Phe Arg Asp Leu
          85           90           95
Val Asp Gly Pro Leu Glu Asp Leu Lys Val Asn Val Trp Glu Gly Arg
          100          105          110
Glu Asp Gly Trp Arg Arg Ser Ser Ala Ala Gly Asp Ala Pro Gln Glu
          115          120          125
Leu Pro Gln Val Glu Ala Ala Glu Arg Gly Arg Lys Leu Asp Asp Ala
          130          135          140
Thr Asp Val Asp Thr Asn Val Gly Thr Glu Glu Gly Phe Glu Glu Gly
          145          150          155          160
Arg Asn

```

<210> 11  
 <211> 220  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>

<221> CDS

<222> (101)..(220)

<223> FRXA02264

<400> 11

```

tccctgacat ccaggttgaa gcaacgtttg atgacggcac caagctcgtc accgtgcaca 60
atcccatccg ataacccttg atgttttttag gagttttgtc atg atc cca ggc gag 115
                                     Met Ile Pro Gly Glu
                                     1           5
tac atc ctg tcc agc gaa tca ctc acc gga aat gtt ggg cgc gag gcc 163
Tyr Ile Leu Ser Ser Glu Ser Leu Thr Gly Asn Val Gly Arg Glu Ala
          10           15           20
aaa acc atc gaa atc atc aac acc ggt gat agg cct gtg cag att ggt 211

```

Lys Thr Ile Glu Ile Ile Asn Thr Gly Asp Arg Pro Val Gln Ile Gly  
                   25                                  30                                  35

tgc cat ttc  
 Ser His Phe  
                   40

220

<210> 12  
 <211> 40  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 12  
 Met Ile Pro Gly Glu Tyr Ile Leu Ser Ser Glu Ser Leu Thr Gly Asn  
       1                                  5                                  10                                  15  
 Val Gly Arg Glu Ala Lys Thr Ile Glu Ile Ile Asn Thr Gly Asp Arg  
                   20                                  25                                  30  
 Pro Val Gln Ile Gly Ser His Phe  
                   35                                  40

<210> 13  
 <211> 1833  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1810)  
 <223> RXN02274

<400> 13  
 acagggtcgaa gctgctgaac gtggccggaa actagatgac gccactgatg tggacacaaa 60  
 tgtgggcaca gaagaaggct ttgaagaagg tcgaaattaa atg agt ttt gag att 115  
   Met Ser Phe Glu Ile  
   1                                  5  
 tcc cgc aag cag tac acc gac ctt tat ggt cca acc gtt ggc gat tca 163  
 Ser Arg Lys Gln Tyr Thr Asp Leu Tyr Gly Pro Thr Val Gly Asp Ser  
                                   10                                  15                                  20  
 gta cgt ctt gct gat act gag ctt ttt ctg tgt gtg gaa aaa gat tac 211  
 Val Arg Leu Ala Asp Thr Glu Leu Phe Leu Cys Val Glu Lys Asp Tyr  
                                   25                                  30                                  35  
 gca gca atc ggc gaa gaa gta gca ttc ggc ggt ggc aag gtc att cgt 259  
 Ala Ala Ile Gly Glu Glu Val Ala Phe Gly Gly Gly Lys Val Ile Arg  
                                   40                                  45                                  50  
 gat ggc atg ggc caa aat ggc acc ttg gtt cgc gat gta gat att ccc 307  
 Asp Gly Met Gly Gln Asn Gly Thr Leu Val Arg Asp Val Asp Ile Pro  
                   55                                  60                                  65  
 gat acc gtc atc acc aac gtc atc gtc ctt gac tat acg ggt gtg tac 355  
 Asp Thr Val Ile Thr Asn Val Ile Val Leu Asp Tyr Thr Gly Val Tyr  
       70                                  75                                  80                                  85

aaa gct gac gtt gcg ctt cga gat ggc aaa atc ttc cga atc gga aag	403
Lys Ala Asp Val Ala Leu Arg Asp Gly Lys Ile Phe Arg Ile Gly Lys	
90 95 100	
gcc gga aac ccg aat gtc atg gaa aac gtc gac atc gtc atc ggc gtt	451
Ala Gly Asn Pro Asn Val Met Glu Asn Val Asp Ile Val Ile Gly Val	
105 110 115	
gcc acc gac atc att gct ggt gaa ggc aaa atc ctt acc gca ggt ggc	499
Ala Thr Asp Ile Ile Ala Gly Glu Gly Lys Ile Leu Thr Ala Gly Gly	
120 125 130	
atc gac acg cac gtg cac ttc ttg ggc aca gac cag gtc aac act gca	547
Ile Asp Thr His Val His Phe Leu Gly Thr Asp Gln Val Asn Thr Ala	
135 140 145	
tta gca tca ggt atc acc acg atg atc ggt gga ggc acc ggc cca agc	595
Leu Ala Ser Gly Ile Thr Thr Met Ile Gly Gly Gly Thr Gly Pro Ser	
150 155 160 165	
cag gcg tcg atg gct aca act gtc acg cca ggt cag tgg aat acc tac	643
Gln Ala Ser Met Ala Thr Thr Val Thr Pro Gly Gln Trp Asn Thr Tyr	
170 175 180	
aac atg ctt agt gct ttt gaa ggc atg ccc atg aac ttt ggc att ttg	691
Asn Met Leu Ser Ala Phe Glu Gly Met Pro Met Asn Phe Gly Ile Leu	
185 190 195	
ggt aaa ggc cat ggt tct tcc aaa tct ccg ctg gct gag cag gtt cgt	739
Gly Lys Gly His Gly Ser Ser Lys Ser Pro Leu Ala Glu Gln Val Arg	
200 205 210	
gcg ggt gca atc ggt ctg aaa att cac gag gac tgg ggt gcc aca cca	787
Ala Gly Ala Ile Gly Leu Lys Ile His Glu Asp Trp Gly Ala Thr Pro	
215 220 225	
tcg tcg atc aac act gcc cta gaa gta gcc gat gac atg gac atc cag	835
Ser Ser Ile Asn Thr Ala Leu Glu Val Ala Asp Asp Met Asp Ile Gln	
230 235 240 245	
gtg gca ctc cac tcc gat acc ttg aat gag gcc ggt ttt gtg gaa gac	883
Val Ala Leu His Ser Asp Thr Leu Asn Glu Ala Gly Phe Val Glu Asp	
250 255 260	
acc att gaa gcc att gcg ggc cga gtc atc cat acc ttc cac acc gaa	931
Thr Ile Glu Ala Ile Ala Gly Arg Val Ile His Thr Phe His Thr Glu	
265 270 275	
ggt gct ggt ggt gga cac gct cct gac cta atc cga gtg gct gct ctg	979
Gly Ala Gly Gly His Ala Pro Asp Leu Ile Arg Val Ala Ala Leu	
280 285 290	
cca aac gtg ttg cct gca tcc acc aac cca acg ctc cca tac acc cga	1027
Pro Asn Val Leu Pro Ala Ser Thr Asn Pro Thr Leu Pro Tyr Thr Arg	
295 300 305	
aac act gtt gaa gag cac ctg gac atg gtg atg gtt gcc cac cac ctc	1075
Asn Thr Val Glu Glu His Leu Asp Met Val Met Val Ala His His Leu	
310 315 320 325	
aac cca gat att cca gaa gac gtg gct ttt gcg gat tcc cga att cgt	1123

Asn	Pro	Asp	Ile	Pro	Glu	Asp	Val	Ala	Phe	Ala	Asp	Ser	Arg	Ile	Arg		
				330					335					340			
gcc	gaa	acg	att	gca	gcc	gaa	gat	gtg	ctt	cac	gat	atg	ggg	atc	ttc	1171	
Ala	Glu	Thr	Ile	Ala	Ala	Glu	Asp	Val	Leu	His	Asp	Met	Gly	Ile	Phe		
			345					350					355				
tct	atc	acc	tct	tcg	gat	tcc	cag	gcg	atg	ggc	cga	gta	gga	gag	acc	1219	
Ser	Ile	Thr	Ser	Ser	Asp	Ser	Gln	Ala	Met	Gly	Arg	Val	Gly	Glu	Thr		
			360				365					370					
atc	acg	cgc	acg	tgg	cag	gtc	gcc	gac	cat	atg	aaa	cgc	acc	cgt	gga	1267	
Ile	Thr	Arg	Thr	Trp	Gln	Val	Ala	Asp	His	Met	Lys	Arg	Thr	Arg	Gly		
	375					380					385						
tca	cta	acg	gga	gat	gct	cca	tac	aac	gac	aac	aac	cgc	ttg	cgt	cga	1315	
Ser	Leu	Thr	Gly	Asp	Ala	Pro	Tyr	Asn	Asp	Asn	Asn	Arg	Leu	Arg	Arg		
390					395					400					405		
ttc	atc	gca	aaa	tac	acc	atc	aac	cct	gcg	att	gcg	cac	ggg	gtg	gat	1363	
Phe	Ile	Ala	Lys	Tyr	Thr	Ile	Asn	Pro	Ala	Ile	Ala	His	Gly	Val	Asp		
			410					415					420				
tat	gtt	gtt	cgt	tca	gtg	gag	gaa	ggc	aag	ttc	gct	gac	ctc	gtg	ctg	1411	
Tyr	Val	Val	Arg	Ser	Val	Glu	Glu	Gly	Lys	Phe	Ala	Asp	Leu	Val	Leu		
			425					430					435				
tgg	gat	cca	aag	ttc	ttt	ggg	gtg	aaa	cct	gat	ctg	gtg	atc	aag	ggg	1459	
Trp	Asp	Pro	Lys	Phe	Phe	Gly	Val	Lys	Pro	Asp	Leu	Val	Ile	Lys	Gly		
		440				445						450					
ggg	ttg	atg	gtc	aat	tcc	ctc	atg	ggg	gat	tcc	aac	ggg	tcc	att	cca	1507	
Gly	Leu	Met	Val	Asn	Ser	Leu	Met	Gly	Asp	Ser	Asn	Gly	Ser	Ile	Pro		
	455					460					465						
act	ccg	cag	ccc	cgc	acc	ctg	cgc	aat	act	tgg	ggg	gcg	ttt	ggc	cag	1555	
Thr	Pro	Gln	Pro	Arg	Thr	Leu	Arg	Asn	Thr	Trp	Gly	Ala	Phe	Gly	Gln		
470					475					480					485		
gca	gtt	tcc	aga	agc	tcc	att	aca	ttc	cta	tcc	cag	gac	gct	atc	gat	1603	
Ala	Val	Ser	Arg	Ser	Ser	Ile	Thr	Phe	Leu	Ser	Gln	Asp	Ala	Ile	Asp		
				490				495					500				
gca	aat	gtt	cct	gat	ctg	ctg	aat	ctg	agg	aag	cag	atc	cgg	ggc	gtt	1651	
Ala	Asn	Val	Pro	Asp	Leu	Leu	Asn	Leu	Arg	Lys	Gln	Ile	Arg	Gly	Val		
			505					510					515				
cga	ggg	gta	agg	aat	ctg	acc	aaa	cga	gac	atg	aaa	ctc	aat	gca	gaa	1699	
Arg	Gly	Val	Arg	Asn	Leu	Thr	Lys	Arg	Asp	Met	Lys	Leu	Asn	Ala	Glu		
		520				525					530						
atg	cct	gat	att	cgt	gtc	gat	cca	gag	acc	tac	cag	gtg	ttt	gtc	aac	1747	
Met	Pro	Asp	Ile	Arg	Val	Asp	Pro	Glu	Thr	Tyr	Gln	Val	Phe	Val	Asn		
	535					540					545						
ggg	gag	ttg	atc	acc	agc	aag	cca	gca	gag	aca	gtg	cca	atg	gca	cgt	1795	
Gly	Glu	Leu	Ile	Thr	Ser	Lys	Pro	Ala	Glu	Thr	Val	Pro	Met	Ala	Arg		
550					555					560				565			
cgc	tac	ttc	ttg	ttc	taatccgcca	acaaggaagg	aag									1833	
Arg	Tyr	Phe	Leu	Phe													

570

&lt;210&gt; 14

&lt;211&gt; 570

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 14

```

Met Ser Phe Glu Ile Ser Arg Lys Gln Tyr Thr Asp Leu Tyr Gly Pro
 1           5           10           15

Thr Val Gly Asp Ser Val Arg Leu Ala Asp Thr Glu Leu Phe Leu Cys
          20           25           30

Val Glu Lys Asp Tyr Ala Ala Ile Gly Glu Glu Val Ala Phe Gly Gly
          35           40           45

Gly Lys Val Ile Arg Asp Gly Met Gly Gln Asn Gly Thr Leu Val Arg
 50           55           60

Asp Val Asp Ile Pro Asp Thr Val Ile Thr Asn Val Ile Val Leu Asp
 65           70           75           80

Tyr Thr Gly Val Tyr Lys Ala Asp Val Ala Leu Arg Asp Gly Lys Ile
          85           90           95

Phe Arg Ile Gly Lys Ala Gly Asn Pro Asn Val Met Glu Asn Val Asp
          100          105          110

Ile Val Ile Gly Val Ala Thr Asp Ile Ile Ala Gly Glu Gly Lys Ile
          115          120          125

Leu Thr Ala Gly Gly Ile Asp Thr His Val His Phe Leu Gly Thr Asp
          130          135          140

Gln Val Asn Thr Ala Leu Ala Ser Gly Ile Thr Thr Met Ile Gly Gly
          145          150          155          160

Gly Thr Gly Pro Ser Gln Ala Ser Met Ala Thr Thr Val Thr Pro Gly
          165          170          175

Gln Trp Asn Thr Tyr Asn Met Leu Ser Ala Phe Glu Gly Met Pro Met
          180          185          190

Asn Phe Gly Ile Leu Gly Lys Gly His Gly Ser Ser Lys Ser Pro Leu
          195          200          205

Ala Glu Gln Val Arg Ala Gly Ala Ile Gly Leu Lys Ile His Glu Asp
          210          215          220

Trp Gly Ala Thr Pro Ser Ser Ile Asn Thr Ala Leu Glu Val Ala Asp
          225          230          235          240

Asp Met Asp Ile Gln Val Ala Leu His Ser Asp Thr Leu Asn Glu Ala
          245          250          255

Gly Phe Val Glu Asp Thr Ile Glu Ala Ile Ala Gly Arg Val Ile His
          260          265          270

Thr Phe His Thr Glu Gly Ala Gly Gly Gly His Ala Pro Asp Leu Ile

```

275	280	285
Arg Val Ala Ala Leu Pro Asn Val Leu Pro Ala Ser Thr Asn Pro Thr 290 295 300		
Leu Pro Tyr Thr Arg Asn Thr Val Glu Glu His Leu Asp Met Val Met 305 310 315 320		
Val Ala His His Leu Asn Pro Asp Ile Pro Glu Asp Val Ala Phe Ala 325 330 335		
Asp Ser Arg Ile Arg Ala Glu Thr Ile Ala Ala Glu Asp Val Leu His 340 345 350		
Asp Met Gly Ile Phe Ser Ile Thr Ser Ser Asp Ser Gln Ala Met Gly 355 360 365		
Arg Val Gly Glu Thr Ile Thr Arg Thr Trp Gln Val Ala Asp His Met 370 375 380		
Lys Arg Thr Arg Gly Ser Leu Thr Gly Asp Ala Pro Tyr Asn Asp Asn 385 390 395 400		
Asn Arg Leu Arg Arg Phe Ile Ala Lys Tyr Thr Ile Asn Pro Ala Ile 405 410 415		
Ala His Gly Val Asp Tyr Val Val Arg Ser Val Glu Glu Gly Lys Phe 420 425 430		
Ala Asp Leu Val Leu Trp Asp Pro Lys Phe Phe Gly Val Lys Pro Asp 435 440 445		
Leu Val Ile Lys Gly Gly Leu Met Val Asn Ser Leu Met Gly Asp Ser 450 455 460		
Asn Gly Ser Ile Pro Thr Pro Gln Pro Arg Thr Leu Arg Asn Thr Trp 465 470 475 480		
Gly Ala Phe Gly Gln Ala Val Ser Arg Ser Ser Ile Thr Phe Leu Ser 485 490 495		
Gln Asp Ala Ile Asp Ala Asn Val Pro Asp Leu Leu Asn Leu Arg Lys 500 505 510		
Gln Ile Arg Gly Val Arg Gly Val Arg Asn Leu Thr Lys Arg Asp Met 515 520 525		
Lys Leu Asn Ala Glu Met Pro Asp Ile Arg Val Asp Pro Glu Thr Tyr 530 535 540		
Gln Val Phe Val Asn Gly Glu Leu Ile Thr Ser Lys Pro Ala Glu Thr 545 550 555 560		
Val Pro Met Ala Arg Arg Tyr Phe Leu Phe 565 570		

<210> 15  
 <211> 1625  
 <212> DNA  
 <213> Corynebacterium glutamicum



&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1602)

&lt;223&gt; FRXA02274

&lt;400&gt; 15

tac gca gca atc ggc gaa gaa gta gca ttc ggc ggt ggc aag gtc att	48
Tyr Ala Ala Ile Gly Glu Glu Val Ala Phe Gly Gly Gly Lys Val Ile	
1 5 10 15	
cgt gat ggc atg ggc caa aat ggc acc ttg gtt cgc gat gta gat att	96
Arg Asp Gly Met Gly Gln Asn Gly Thr Leu Val Arg Asp Val Asp Ile	
20 25 30	
ccc gat acc gtc atc acc aac gtc atc gtc ctt gac tat acg ggt gtg	144
Pro Asp Thr Val Ile Thr Asn Val Ile Val Leu Asp Tyr Thr Gly Val	
35 40 45	
tac aaa gct gac gtt gcg ctt cga gat ggc aaa atc ttc cga atc gga	192
Tyr Lys Ala Asp Val Ala Leu Arg Asp Gly Lys Ile Phe Arg Ile Gly	
50 55 60	
aag gcc gga aac ccg aat gtc atg gaa aac gtc gac atc gtc atc ggc	240
Lys Ala Gly Asn Pro Asn Val Met Glu Asn Val Asp Ile Val Ile Gly	
65 70 75 80	
gtt gcc acc gac atc att gct ggt gaa ggc aaa atc ctt acc gca ggt	288
Val Ala Thr Asp Ile Ile Ala Gly Glu Gly Lys Ile Leu Thr Ala Gly	
85 90 95	
ggc atc gac acg cac gtg cac ttc ttg ggc aca gac cag gtc aac act	336
Gly Ile Asp Thr His Val His Phe Leu Gly Thr Asp Gln Val Asn Thr	
100 105 110	
gca tta gca tca ggt atc acc acg atg atc ggt gga ggc acc ggc cca	384
Ala Leu Ala Ser Gly Ile Thr Thr Met Ile Gly Gly Thr Gly Pro	
115 120 125	
agc cag gcg tcg atg gct aca act gtc acg cca ggt cag tgg aat acc	432
Ser Gln Ala Ser Met Ala Thr Thr Val Thr Pro Gly Gln Trp Asn Thr	
130 135 140	
tac aac atg ctt agt gct ttt gaa ggc atg ccc atg aac ttt ggc att	480
Tyr Asn Met Leu Ser Ala Phe Glu Gly Met Pro Met Asn Phe Gly Ile	
145 150 155 160	
ttg ggt aaa ggc cat ggt tct tcc aaa tct ccg ctg gct gag cag gtt	528
Leu Gly Lys Gly His Gly Ser Ser Lys Ser Pro Leu Ala Glu Gln Val	
165 170 175	
cgt gcg ggt gca atc ggt ctg aaa att cac gag gac tgg ggt gcc aca	576
Arg Ala Gly Ala Ile Gly Leu Lys Ile His Glu Asp Trp Gly Ala Thr	
180 185 190	
cca tcg tcg atc aac act gcc cta gaa gta gcc gat gac atg gac atc	624
Pro Ser Ser Ile Asn Thr Ala Leu Glu Val Ala Asp Asp Met Asp Ile	
195 200 205	
cag gtg gca ctc cac tcc gat acc ttg aat gag gcc ggt ttt gtg gaa	672
Gln Val Ala Leu His Ser Asp Thr Leu Asn Glu Ala Gly Phe Val Glu	

210	215	220	
gac acc att gaa gcc att gcg ggc cga gtc atc cat acc ttc cac acc Asp Thr Ile Glu Ala Ile Ala Gly Arg Val Ile His Thr Phe His Thr 225 230 235 240			720
gaa ggt gct ggt ggt gga cac gct cct gac cta atc cga gtg gct gct Glu Gly Ala Gly Gly Gly His Ala Pro Asp Leu Ile Arg Val Ala Ala 245 250 255			768
ctg cca aac gtg ttg cct gca tcc acc aac cca acg ctc cca tac acc Leu Pro Asn Val Leu Pro Ala Ser Thr Asn Pro Thr Leu Pro Tyr Thr 260 265 270			816
cga aac act gtt gaa gag cac ctg gac atg gtg atg gtt gcc cac cac Arg Asn Thr Val Glu Glu His Leu Asp Met Val Met Val Ala His His 275 280 285			864
ctc aac cca gat att cca gaa gac gtg gct ttt gcg gat tcc cga att Leu Asn Pro Asp Ile Pro Glu Asp Val Ala Phe Ala Asp Ser Arg Ile 290 295 300			912
cgt gcc gaa acg att gca gcc gaa gat gtg ctt cac gat atg ggt atc Arg Ala Glu Thr Ile Ala Ala Glu Asp Val Leu His Asp Met Gly Ile 305 310 315 320			960
ttc tct atc acc tct tcg gat tcc cag gcg atg ggc cga gta gga gag Phe Ser Ile Thr Ser Ser Asp Ser Gln Ala Met Gly Arg Val Gly Glu 325 330 335			1008
acc atc acg cgc acg tgg cag gtc gcc gac cat atg aaa cgc acc cgt Thr Ile Thr Arg Thr Trp Gln Val Ala Asp His Met Lys Arg Thr Arg 340 345 350			1056
gga tca cta acg gga gat gct cca tac aac gac aac aac cgc ttg cgt Gly Ser Leu Thr Gly Asp Ala Pro Tyr Asn Asp Asn Asn Arg Leu Arg 355 360 365			1104
cga ttc atc gca aaa tac acc atc aac cct gcg att gcg cac ggt gtg Arg Phe Ile Ala Lys Tyr Thr Ile Asn Pro Ala Ile Ala His Gly Val 370 375 380			1152
gat tat gtt gtt cgt tca gtg gag gaa ggc aag ttc gct gac ctc gtg Asp Tyr Val Val Arg Ser Val Glu Glu Gly Lys Phe Ala Asp Leu Val 385 390 395 400			1200
ctg tgg gat cca aag ttc ttt ggt gtg aaa cct gat ctg gtg atc aag Leu Trp Asp Pro Lys Phe Phe Gly Val Lys Pro Asp Leu Val Ile Lys 405 410 415			1248
ggt ggg ttg atg gtc aat tcc ctc atg ggt gat tcc aac ggt tcc att Gly Gly Leu Met Val Asn Ser Leu Met Gly Asp Ser Asn Gly Ser Ile 420 425 430			1296
cca act ccg cag ccc cgc acc ctg cgc aat act tgg ggt gcg ttt ggc Pro Thr Pro Gln Pro Arg Thr Leu Arg Asn Thr Trp Gly Ala Phe Gly 435 440 445			1344
cag gca gtt tcc aga agc tcc att aca ttc cta tcc cag gac gct atc Gln Ala Val Ser Arg Ser Ser Ile Thr Phe Leu Ser Gln Asp Ala Ile 450 455 460			1392

gat gca aat gtt cct gat ctg ctg aat ctg agg aag cag atc cgg ggc 1440  
 Asp Ala Asn Val Pro Asp Leu Leu Asn Leu Arg Lys Gln Ile Arg Gly  
 465 470 475 480  
 gtt cga ggt gta agg aat ctg acc aaa cga gac atg aaa ctc aat gca 1488  
 Val Arg Gly Val Arg Asn Leu Thr Lys Arg Asp Met Lys Leu Asn Ala  
 485 490 495  
 gaa atg cct gat att cgt gtc gat cca gag acc tac cag gtg ttt gtc 1536  
 Glu Met Pro Asp Ile Arg Val Asp Pro Glu Thr Tyr Gln Val Phe Val  
 500 505 510  
 aac ggt gag ttg atc acc agc aag cca gca gag aca gtg cca atg gca 1584  
 Asn Gly Glu Leu Ile Thr Ser Lys Pro Ala Glu Thr Val Pro Met Ala  
 515 520 525  
 cgt cgc tac ttc ttg ttc taatccgcca acaaggaagg aag 1625  
 Arg Arg Tyr Phe Leu Phe  
 530

<210> 16  
 <211> 534  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 16  
 Tyr Ala Ala Ile Gly Glu Glu Val Ala Phe Gly Gly Gly Lys Val Ile  
 1 5 10 15  
 Arg Asp Gly Met Gly Gln Asn Gly Thr Leu Val Arg Asp Val Asp Ile  
 20 25 30  
 Pro Asp Thr Val Ile Thr Asn Val Ile Val Leu Asp Tyr Thr Gly Val  
 35 40 45  
 Tyr Lys Ala Asp Val Ala Leu Arg Asp Gly Lys Ile Phe Arg Ile Gly  
 50 55 60  
 Lys Ala Gly Asn Pro Asn Val Met Glu Asn Val Asp Ile Val Ile Gly  
 65 70 75 80  
 Val Ala Thr Asp Ile Ile Ala Gly Glu Gly Lys Ile Leu Thr Ala Gly  
 85 90 95  
 Gly Ile Asp Thr His Val His Phe Leu Gly Thr Asp Gln Val Asn Thr  
 100 105 110  
 Ala Leu Ala Ser Gly Ile Thr Thr Met Ile Gly Gly Gly Thr Gly Pro  
 115 120 125  
 Ser Gln Ala Ser Met Ala Thr Thr Val Thr Pro Gly Gln Trp Asn Thr  
 130 135 140  
 Tyr Asn Met Leu Ser Ala Phe Glu Gly Met Pro Met Asn Phe Gly Ile  
 145 150 155 160  
 Leu Gly Lys Gly His Gly Ser Ser Lys Ser Pro Leu Ala Glu Gln Val  
 165 170 175

Arg Ala Gly Ala Ile Gly Leu Lys Ile His Glu Asp Trp Gly Ala Thr  
 180 185 190  
 Pro Ser Ser Ile Asn Thr Ala Leu Glu Val Ala Asp Asp Met Asp Ile  
 195 200 205  
 Gln Val Ala Leu His Ser Asp Thr Leu Asn Glu Ala Gly Phe Val Glu  
 210 215 220  
 Asp Thr Ile Glu Ala Ile Ala Gly Arg Val Ile His Thr Phe His Thr  
 225 230 235 240  
 Glu Gly Ala Gly Gly Gly His Ala Pro Asp Leu Ile Arg Val Ala Ala  
 245 250 255  
 Leu Pro Asn Val Leu Pro Ala Ser Thr Asn Pro Thr Leu Pro Tyr Thr  
 260 265 270  
 Arg Asn Thr Val Glu Glu His Leu Asp Met Val Met Val Ala His His  
 275 280 285  
 Leu Asn Pro Asp Ile Pro Glu Asp Val Ala Phe Ala Asp Ser Arg Ile  
 290 295 300  
 Arg Ala Glu Thr Ile Ala Ala Glu Asp Val Leu His Asp Met Gly Ile  
 305 310 315 320  
 Phe Ser Ile Thr Ser Ser Asp Ser Gln Ala Met Gly Arg Val Gly Glu  
 325 330 335  
 Thr Ile Thr Arg Thr Trp Gln Val Ala Asp His Met Lys Arg Thr Arg  
 340 345 350  
 Gly Ser Leu Thr Gly Asp Ala Pro Tyr Asn Asp Asn Asn Arg Leu Arg  
 355 360 365  
 Arg Phe Ile Ala Lys Tyr Thr Ile Asn Pro Ala Ile Ala His Gly Val  
 370 375 380  
 Asp Tyr Val Val Arg Ser Val Glu Glu Gly Lys Phe Ala Asp Leu Val  
 385 390 395 400  
 Leu Trp Asp Pro Lys Phe Phe Gly Val Lys Pro Asp Leu Val Ile Lys  
 405 410 415  
 Gly Gly Leu Met Val Asn Ser Leu Met Gly Asp Ser Asn Gly Ser Ile  
 420 425 430  
 Pro Thr Pro Gln Pro Arg Thr Leu Arg Asn Thr Trp Gly Ala Phe Gly  
 435 440 445  
 Gln Ala Val Ser Arg Ser Ser Ile Thr Phe Leu Ser Gln Asp Ala Ile  
 450 455 460  
 Asp Ala Asn Val Pro Asp Leu Leu Asn Leu Arg Lys Gln Ile Arg Gly  
 465 470 475 480  
 Val Arg Gly Val Arg Asn Leu Thr Lys Arg Asp Met Lys Leu Asn Ala  
 485 490 495  
 Glu Met Pro Asp Ile Arg Val Asp Pro Glu Thr Tyr Gln Val Phe Val

<400> 18  
Leu His Ile Thr Pro Arg Glu Gln Glu Lys Leu Met Ile Val Val Ala  
1 5 10 15

<400> 19															
accattctgc ctgactaatc tgcgcaccga tgatggtttg gataagggtct tggaatggat															60
ccgccatgag gtgatgatgc aggacttgca ggaagcctaa atg aca caa acc caa															115
Met Thr Gln Thr Gln															
1 5															
cca gtg gga acc ctg cga ctg acc atc gat gat caa gga ccc caa ggt															163
Pro Val Gly Thr Leu Arg Leu Thr Ile Asp Asp Gln Gly Pro Gln Gly															
10 15 20															
caa agc cgt gcg gtg gag caa ttt cac cag ggt gcg ctt cga gtc atc															211
Gln Ser Arg Ala Val Glu Gln Phe His Gln Gly Ala Leu Arg Val Ile															
25 30 35															
cgg cca cac tac ttg gat gat tcc gga cag gtt tgc tac acc atc att															259
Arg Pro His Tyr Leu Asp Asp Ser Gly Gln Val Cys Tyr Thr Ile Ile															
40 45 50															
gcc att ggt ggc gga tac ctg ggc ggc gat gtg tat gag cag caa ttc															307
Ala Ile Gly Gly Gly Tyr Leu Gly Gly Asp Val Tyr Gln Gln Phe															
55 60 65															
acg atc aaa gac aac gca aaa gct ttg atc acc acg caa tcg gcc acc															355
Thr Ile Lys Asp Asn Ala Lys Ala Leu Ile Thr Thr Gln Ser Ala Thr															
70 75 80 85															
aag att tat cgc aca ccg caa gga cca gcc acg cag cac acc gaa atc															403
Lys Ile Tyr Arg Thr Pro Gln Gly Pro Ala Thr Gln His Thr Glu Ile															
90 95 100															

aac gtc ggt gaa aat gct gtg ctg gaa tac ttg gcg gat caa acc atc 451  
 Asn Val Gly Glu Asn Ala Val Leu Glu Tyr Leu Ala Asp Gln Thr Ile  
 105 110 115  
  
 gcg tac cgg gag gcc acc tat cat caa ttc acc aag gtg gcg ctg cac 499  
 Ala Tyr Arg Glu Ala Thr Tyr His Gln Phe Thr Lys Val Ala Leu His  
 120 125 130  
  
 ccg agc gca acg ttt gtg atg agc gaa caa atc acc cca ggc tgg cac 547  
 Pro Ser Ala Thr Phe Val Met Ser Glu Gln Ile Thr Pro Gly Trp His  
 135 140 145  
  
 ccc gac ggc aaa cac ttt gct tac gat gaa atg cgt cta cac acc gaa 595  
 Pro Asp Gly Lys His Phe Ala Tyr Asp Glu Met Arg Leu His Thr Glu  
 150 155 160 165  
  
 atc acg gac tcc acc aca ggg cga ctc gtg ctc ttg gat aat tta ctg 643  
 Ile Thr Asp Ser Thr Thr Gly Arg Leu Val Leu Leu Asp Asn Leu Leu  
 170 175 180  
  
 ctc cgg ccg gac tcc cga gag gga agt ttt ggg tgg acg gaa cag tac 691  
 Leu Arg Pro Asp Ser Arg Glu Gly Ser Phe Gly Trp Thr Glu Gln Tyr  
 185 190 195  
  
 aca cat tca ggg cag atg att gtg atg ggg gaa ggc gtc gat aag cag 739  
 Thr His Ser Gly Gln Met Ile Val Met Gly Glu Gly Val Asp Lys Gln  
 200 205 210  
  
 ctt gtt gct gag ctg aat gag caa ctt gcc gcg cac cct gat gtg tac 787  
 Leu Val Ala Glu Leu Asn Glu Gln Leu Ala Ala His Pro Asp Val Tyr  
 215 220 225  
  
 ggc gcc gtc aat ttc tta agc gcg ccg ggc acg tta ctg cgc gga ttt 835  
 Gly Ala Val Asn Phe Leu Ser Ala Pro Gly Thr Leu Leu Arg Gly Phe  
 230 235 240 245  
  
 att gcg cgc acg ctg agc aac cgc act gag gag ttg att aac ctg cac 883  
 Ile Ala Arg Thr Leu Ser Asn Arg Thr Glu Glu Leu Ile Asn Leu His  
 250 255 260  
  
 gaa cac att gcg tcg ctg ttg cgc ggg ccg tgg cgc ggg cag gaa ccg 931  
 Glu His Ile Ala Ser Leu Leu Arg Gly Arg Trp Arg Gly Gln Glu Pro  
 265 270 275  
  
 gtg aat ttg ccg aag tac tagacggcgt cgagaaatcg aag 972  
 Val Asn Leu Arg Lys Tyr  
 280

&lt;210&gt; 20

&lt;211&gt; 283

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 20

Met Thr Gln Thr Gln Pro Val Gly Thr Leu Arg Leu Thr Ile Asp Asp  
 1 5 10 15

Gln Gly Pro Gln Gly Gln Ser Arg Ala Val Glu Gln Phe His Gln Gly  
 20 25 30

Ala Leu Arg Val Ile Arg Pro His Tyr Leu Asp Asp Ser Gly Gln Val  
           35                                  40                                  45  
 Cys Tyr Thr Ile Ile Ala Ile Gly Gly Gly Tyr Leu Gly Gly Asp Val  
           50                                  55                                  60  
 Tyr Glu Gln Gln Phe Thr Ile Lys Asp Asn Ala Lys Ala Leu Ile Thr  
           65                                  70                                  75                                  80  
 Thr Gln Ser Ala Thr Lys Ile Tyr Arg Thr Pro Gln Gly Pro Ala Thr  
                                   85                                  90                                  95  
 Gln His Thr Glu Ile Asn Val Gly Glu Asn Ala Val Leu Glu Tyr Leu  
                                   100                                  105                                  110  
 Ala Asp Gln Thr Ile Ala Tyr Arg Glu Ala Thr Tyr His Gln Phe Thr  
                                   115                                  120                                  125  
 Lys Val Ala Leu His Pro Ser Ala Thr Phe Val Met Ser Glu Gln Ile  
           130                                  135                                  140  
 Thr Pro Gly Trp His Pro Asp Gly Lys His Phe Ala Tyr Asp Glu Met  
           145                                  150                                  155                                  160  
 Arg Leu His Thr Glu Ile Thr Asp Ser Thr Thr Gly Arg Leu Val Leu  
                                   165                                  170                                  175  
 Leu Asp Asn Leu Leu Leu Arg Pro Asp Ser Arg Glu Gly Ser Phe Gly  
                                   180                                  185                                  190  
 Trp Thr Glu Gln Tyr Thr His Ser Gly Gln Met Ile Val Met Gly Glu  
           195                                  200                                  205  
 Gly Val Asp Lys Gln Leu Val Ala Glu Leu Asn Glu Gln Leu Ala Ala  
           210                                  215                                  220  
 His Pro Asp Val Tyr Gly Ala Val Asn Phe Leu Ser Ala Pro Gly Thr  
           225                                  230                                  235                                  240  
 Leu Leu Arg Gly Phe Ile Ala Arg Thr Leu Ser Asn Arg Thr Glu Glu  
                                   245                                  250                                  255  
 Leu Ile Asn Leu His Glu His Ile Ala Ser Leu Leu Arg Gly Arg Trp  
                                   260                                  265                                  270  
 Arg Gly Gln Glu Pro Val Asn Leu Arg Lys Tyr  
           275                                  280

&lt;210&gt; 21

&lt;211&gt; 594

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (101)..(571)

&lt;223&gt; RXA02275

&lt;400&gt; 21

gtttgtcaac ggtgagttga tcaccagcaa gccagcagag acagtgccaa tggcacgtcg 60



ctactttcttg ttctaataccg ccaacaagga aggaagatcc atg att atc act gcg 115  
Met Ile Ile Thr Ala  
1 5

atc gac acc aac atc tac gat gaa ccg gag ttt gtt gaa gga cgc gat 163  
Ile Asp Thr Asn Ile Tyr Asp Glu Pro Glu Phe Val Glu Gly Arg Asp  
10 15 20

gtc atc ggt gtg cgc ttt gaa gat tta gtt ttg gat aag cgc att caa 211  
Val Ile Gly Val Arg Phe Glu Asp Leu Val Leu Asp Lys Arg Ile Gln  
25 30 35

cgg gtt gca ctc ccc gga gga gaa gaa ctg ggg ttg cgg tta aac cac 259  
Arg Val Ala Leu Pro Gly Gly Glu Glu Leu Gly Leu Arg Leu Asn His  
40 45 50

ggg cat ccg att ctg cgt gaa ggt gat gtg ttg aaa gct gat gat aag 307  
Gly His Pro Ile Leu Arg Glu Gly Asp Val Leu Lys Ala Asp Asp Lys  
55 60 65

acg gta ttt gtg gtg gag att atc ccc acg gat gtt tta gtt atc acg 355  
Thr Val Phe Val Val Glu Ile Ile Pro Thr Asp Val Leu Val Ile Thr  
70 75 80 85

cca agc gat att cac cag atg gga ttt gtg gcg cac tcc ctg gga aac 403  
Pro Ser Asp Ile His Gln Met Gly Phe Val Ala His Ser Leu Gly Asn  
90 95 100

agg cac ctg cca gca cag ttt tcc aag cca ggt gaa ttg aca gag aag 451  
Arg His Leu Pro Ala Gln Phe Ser Lys Pro Gly Glu Leu Thr Glu Lys  
105 110 115

gca gcc atg atc gtg caa tac gat cac acg gtg gtc agc ttc ttg gat 499  
Ala Ala Met Ile Val Gln Tyr Asp His Thr Val Val Ser Phe Leu Asp  
120 125 130

gag cac ggc atc gag tat cag cgc acc gaa ctt gtt ccg cca att cct 547  
Glu His Gly Ile Glu Tyr Gln Arg Thr Glu Leu Val Pro Pro Ile Pro  
135 140 145

ttc agg cat agc ggg cac aca cat tgatggatct tgacgctgat ttt 594  
Phe Arg His Ser Gly His Thr His  
150 155

&lt;210&gt; 22

&lt;211&gt; 157

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 22

Met Ile Ile Thr Ala Ile Asp Thr Asn Ile Tyr Asp Glu Pro Glu Phe  
1 5 10 15

Val Glu Gly Arg Asp Val Ile Gly Val Arg Phe Glu Asp Leu Val Leu  
20 25 30

Asp Lys Arg Ile Gln Arg Val Ala Leu Pro Gly Gly Glu Glu Leu Gly  
35 40 45

Leu Arg Leu Asn His Gly His Pro Ile Leu Arg Glu Gly Asp Val Leu  
 50 55 60  
 Lys Ala Asp Asp Lys Thr Val Phe Val Val Glu Ile Ile Pro Thr Asp  
 65 70 75 80  
 Val Leu Val Ile Thr Pro Ser Asp Ile His Gln Met Gly Phe Val Ala  
 85 90 95  
 His Ser Leu Gly Asn Arg His Leu Pro Ala Gln Phe Ser Lys Pro Gly  
 100 105 110  
 Glu Leu Thr Glu Lys Ala Ala Met Ile Val Gln Tyr Asp His Thr Val  
 115 120 125  
 Val Ser Phe Leu Asp Glu His Gly Ile Glu Tyr Gln Arg Thr Glu Leu  
 130 135 140  
 Val Pro Pro Ile Pro Phe Arg His Ser Gly His Thr His  
 145 150 155

&lt;210&gt; 23

&lt;211&gt; 801

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (101)..(778)

&lt;223&gt; RXA02276

&lt;400&gt; 23

atcacacggt ggtcagcttc ttggatgagc acggcatcga gtatcagcgc accgaacttg 60  
 ttccgccaat tcctttcagg catagcgggc acacacattg atg gat ctt gac gct 115  
 Met Asp Leu Asp Ala 5  
 gat ttt ctg ctg ttg cat tta tgg gat tca gca ctt cca acg gga gcg 163  
 Asp Phe Leu Leu Leu His Leu Ser Asp Ser Ala Leu Pro Thr Gly Ala 20  
 10 15  
 ttt gcg cac tca ttt gga ttt gaa act tat atg gat gca gag cga atc 211  
 Phe Ala His Ser Phe Gly Phe Glu Thr Tyr Met Asp Ala Glu Arg Ile 35  
 25 30  
 acc aat gca gag gag ttc caa gac tgg ctg aaa gtc ctg ctt aag gtg 259  
 Thr Asn Ala Glu Glu Phe Gln Asp Trp Leu Lys Val Leu Leu Lys Val 50  
 40 45  
 caa ttg acc agc tct gat gct ttg gca atg agg atg ttt tac gcc acc 307  
 Gln Leu Thr Ser Ser Asp Ala Leu Ala Met Arg Met Phe Tyr Ala Thr 65  
 55 60  
 ccg acg gtg tct gag ctg aaa cgg ctg gat gag cgc ctt ttt gct gga 355  
 Pro Thr Val Ser Glu Leu Lys Arg Leu Asp Glu Arg Leu Phe Ala Gly 85  
 70 75 80  
 act ccg gcg aga gaa att cgg gaa gct aat gct cga atg ggt acg cgc 403  
 Thr Pro Ala Arg Glu Ile Arg Glu Ala Asn Ala Arg Met Gly Thr Arg

	90	95	100	
atg gca gag atc gtg gct gaa acc tac tcc gtg ccc ctg att gtt gag				451
Met Ala Glu Ile Val Ala Glu Thr Tyr Ser Val Pro Leu Ile Val Glu				
	105	110	115	
tat ctc gaa ttg att caa cat cga gag cta tca ggg cac ccg gct ttg				499
Tyr Leu Glu Leu Ile Gln His Arg Glu Leu Ser Gly His Pro Ala Leu				
	120	125	130	
gct ttg gct ctt gcc acc cac agc gcg ggg att gat gtg gat cga gca				547
Ala Leu Ala Leu Ala Thr His Ser Ala Gly Ile Asp Val Asp Arg Ala				
	135	140	145	
atc cac gct cac ctc acg gca acg gtg agt tcg ctg atc caa aat gcg				595
Ile His Ala His Leu Thr Ala Thr Val Ser Ser Leu Ile Gln Asn Ala				
	150	155	160	165
gtt cgt ggc atc cca ctg ggg caa atg gca ggt cag cgg gtg atg ttc				643
Val Arg Gly Ile Pro Leu Gly Gln Met Ala Gly Gln Arg Val Met Phe				
	170	175	180	
gcc atg cgt gag cat atc ggt gcg gcc gtg aaa cgt agc gcg aac ttg				691
Ala Met Arg Glu His Ile Gly Ala Ala Val Lys Arg Ser Ala Asn Leu				
	185	190	195	
gat gag att gat ttc tgt tcg ggt gat cca ggc ttg gat att tca caa				739
Asp Glu Ile Asp Phe Cys Ser Gly Asp Pro Gly Leu Asp Ile Ser Gln				
	200	205	210	
atg gtt cat gaa acc caa cgc gca cga cta ttt atg agt taagaaggag				788
Met Val His Glu Thr Gln Arg Ala Arg Leu Phe Met Ser				
	215	220	225	
aaaagaaaca tgg				801
<210> 24				
<211> 226				
<212> PRT				
<213> Corynebacterium glutamicum				
<400> 24				
Met Asp Leu Asp Ala Asp Phe Leu Leu Leu His Leu Ser Asp Ser Ala				
1 5 10 15				
Leu Pro Thr Gly Ala Phe Ala His Ser Phe Gly Phe Glu Thr Tyr Met				
20 25 30				
Asp Ala Glu Arg Ile Thr Asn Ala Glu Glu Phe Gln Asp Trp Leu Lys				
35 40 45				
Val Leu Leu Lys Val Gln Leu Thr Ser Ser Asp Ala Leu Ala Met Arg				
50 55 60				
Met Phe Tyr Ala Thr Pro Thr Val Ser Glu Leu Lys Arg Leu Asp Glu				
65 70 75 80				
Arg Leu Phe Ala Gly Thr Pro Ala Arg Glu Ile Arg Glu Ala Asn Ala				
85 90 95				

Gly Arg Glu Ile Ala Ala Lys Tyr Ala Glu Ala Ile Tyr Ser Val Ala  
 200 205 210

tgg gat ttg gag caa gcg caa gat tat cgc tct gat att cat gct cgt 787  
 Trp Asp Leu Glu Gln Ala Gln Asp Tyr Arg Ser Asp Ile His Ala Arg  
 215 220 225

gcc act gcc cag ggt cgc gag ccc atg ccg gtg ctt cct ggt ttg gtg 835  
 Ala Thr Ala Gln Gly Arg Glu Pro Met Pro Val Leu Pro Gly Leu Val  
 230 235 240 245

act ttt gtt ggc acg acc gtg gaa gaa gcg cgt gca aaa cag cag gct 883  
 Thr Phe Val Gly Thr Thr Val Glu Glu Ala Arg Ala Lys Gln Gln Ala  
 250 255 260

ctt aat gcg ttg ctg ccg gtc aaa gac tca cta aat cag ttg agt ttc 931  
 Leu Asn Ala Leu Leu Pro Val Lys Asp Ser Leu Asn Gln Leu Ser Phe  
 265 270 275

ttt gtg ggt caa gat tgc tcg acg tgg gat ttg gat gca cct ccc cca 979  
 Phe Val Gly Gln Asp Cys Ser Thr Trp Asp Leu Asp Ala Pro Pro Pro  
 280 285 290

cca ctg cca ccg cta gaa gag ttt tcc ggt cct aaa ggc agg tac gaa 1027  
 Pro Leu Pro Pro Leu Glu Glu Phe Ser Gly Pro Lys Gly Arg Tyr Glu  
 295 300 305

acg gtc ctg cgg  
 Thr Val Leu Arg 1039  
 310

<210> 440  
 <211> 313  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 440  
 Val Glu Gly Ser Val Glu Lys Leu Gly Leu Ile Ser Trp Trp Glu Glu  
 1 5 10 15

Leu Ala Arg Thr Ala Glu Arg Gly Lys Leu Asp Ala Val Phe Leu Ala  
 20 25 30

Asp Gly Gln Ala Ile Asn Pro Val Gly Leu Glu Asn Gly Pro Gly Trp  
 35 40 45

Phe Leu Glu Pro Val Thr Ala Leu Thr Ala Met Ala Arg Ala Thr Asn  
 50 55 60

Asn Ile Gly Leu Ile Ser Thr Ile Ser Ser Thr Phe Trp Gln Pro Phe  
 65 70 75 80

His Ala Ala Arg Met Ile Ala Ser Leu Asp His Ile Ser Gly Gly Arg  
 85 90 95

Ala Gly Ile Asn Val Val Thr Ser Met Thr Asp Ala Glu Ala Arg Asn  
 100 105 110

His Gly Met Asp Ala Leu Pro Gly His Asp Val Arg Tyr Ala Arg Ala  
 115 120 125

Ala Glu Phe Ile Glu Thr Ile Thr Ala Leu Trp Asp Ser Trp Pro Ala  
 130 135 140

Glu Ser Leu Val Met Asp Arg Ala Gly Lys Phe Ala Asp Ser Ser Leu  
 145 150 155 160

Ile Lys Ser Ile Asp His Asp Gly Glu Phe Phe Gln Val Ala Gly Pro  
 165 170 175

Leu Asn Ile Pro Ser Pro Pro Gln Gly Arg Pro Val Leu Phe Gln Ala  
 180 185 190

Gly Ser Ser Pro Gln Gly Arg Glu Ile Ala Ala Lys Tyr Ala Glu Ala  
 195 200 205

BGI-128CPVC  
 Phe Tyr Ser Val Ala Trp Asp Leu Glu Gln Ala Gln Asp Tyr Arg Ser  
 210 215 220

Asp Ile His Ala Arg Ala Thr Ala Gln Gly Arg Glu Pro Met Pro Val  
 225 230 235 240

Leu Pro Gly Leu Val Thr Phe Val Gly Thr Thr Val Glu Glu Ala Arg  
 245 250 255

Ala Lys Gln Gln Ala Leu Asn Ala Leu Leu Pro Val Lys Asp Ser Leu  
 260 265 270

Asn Gln Leu Ser Phe Phe Val Gly Gln Asp Cys Ser Thr Trp Asp Leu  
 275 280 285

Asp Ala Pro Pro Pro Pro Leu Pro Pro Leu Glu Glu Phe Ser Gly Pro  
 290 295 300

Lys Gly Arg Tyr Glu Thr Val Leu Arg  
 305 310

BGI-128CPVC 215

2

Ile Phe Val Glu Ser Gly Gly Asp Asn Leu Ser Ala Thr Phe Ser Pro  
 100 105 110

Glu Leu Val Asp Phe Ser Ile Tyr Ile Ile Asp Val Ala Gln Gly Glu  
 115 120 125

Lys Ile Pro Arg Lys Ala Gly Gln Gly Met Ile Lys Ser Asp Leu Phe  
 130 135 140

Ile Ile Asn Lys Thr Asp Leu Ala Pro Tyr Val Gly Ala Asn Leu Asp  
 145 150 155 160

Val Met Val Glu Asp Ala Lys Ala Phe Arg Lys Asn Lys Pro Phe Cys  
 165 170 175

Leu Thr Asn Leu Arg Thr Asp Asp Gly Leu Asp Lys Val Leu Glu Trp  
 180 185 190

Ile Arg His Glu Val Met Met Gln Asp Leu Gln Glu Ala  
 195 200 205

<210> 27  
 <211> 1119  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1096)  
 <223> RXA02603

<400> 27  
 gcatggtgct tggcgtgcat tacccaaccg atgtgctagc cggcgcgttg ttgggagcag 60

cgaccgcaga ggccgtccat aagatcgaaa gggctacgaa gtg agc gaa cac gcc 115  
 Val Ser Glu His Ala  
 1 5

gct gaa cat cac cgc gat acc caa aat ttc tta acc tcc gaa ccg cac 163  
 Ala Glu His His Arg Asp Thr Gln Asn Phe Leu Thr Ser Glu Pro His  
 10 15 20

acc acg gca atc gaa gac aac aag aag cgc caa ccg ccg aaa aac ctt 211  
 Thr Thr Ala Ile Glu Asp Asn Lys Lys Arg Gln Pro Pro Lys Asn Leu  
 25 30 35

gct gac ggc atg atc aag gcg ctg cgc ccc aag cag tgg gtc aag aac 259  
 Ala Asp Gly Met Ile Lys Ala Leu Arg Pro Lys Gln Trp Val Lys Asn  
 40 45 50

gtt ctt gtg cta gca gca cca ctt gct gct ggt gca gat gcg atc ttc 307  
 Val Leu Val Leu Ala Ala Pro Leu Ala Ala Gly Ala Asp Ala Ile Phe  
 55 60 65

aac cag cgc acg atc atc gac gtt gct atc gca ttc gta gtg ttc tgc 355  
 Asn Gln Arg Thr Ile Ile Asp Val Ala Ile Ala Phe Val Val Phe Cys  
 70 75 80 85

ttc ggt gca tca gcc att tac ttg gtt aat gat gcc cgt gac gtg gaa 403  
 Phe Gly Ala Ser Ala Ile Tyr Leu Val Asn Asp Ala Arg Asp Val Glu

90								95				100				
gct	gac	cgc	gag	cac	cca	acc	aag	cgt	ttc	cgc	ccc	atc	gct	gca	gga	451
Ala	Asp	Arg	Glu	His	Pro	Thr	Lys	Arg	Phe	Arg	Pro	Ile	Ala	Ala	Gly	
105								110				115				
gtc	ctg	cca	gta	gga	atg	gca	tac	ggc	atg	gcc	gtg	gcg	ctc	att	gca	499
Val	Leu	Pro	Val	Gly	Met	Ala	Tyr	Gly	Met	Ala	Val	Ala	Leu	Ile	Ala	
120								125				130				
cta	tcc	atc	gga	ctg	tct	ttc	ctc	gcc	acc	gac	ggc	gtg	gca	ctt	gcc	547
Leu	Ser	Ile	Gly	Leu	Ser	Phe	Leu	Ala	Thr	Asp	Gly	Val	Ala	Leu	Ala	
135								140				145				
tgc	gtg	att	ggc	gtg	tac	att	gcg	ctg	cag	ctg	gga	tac	tgc	ttc	ggt	595
Cys	Val	Ile	Gly	Val	Tyr	Ile	Ala	Leu	Gln	Leu	Gly	Tyr	Cys	Phe	Gly	
150								155				160				165
tgg	aag	cac	atg	cca	gtg	atc	gat	att	gcg	ctt	gtc	tcc	tcc	gga	ttc	643
Trp	Lys	His	Met	Pro	Val	Ile	Asp	Ile	Ala	Leu	Val	Ser	Ser	Gly	Phe	
170								175				180				
atg	ctc	cgc	gca	atg	gca	ggc	ggc	gtc	gca	gca	ggc	atc	gag	cta	tcc	691
Met	Leu	Arg	Ala	Met	Ala	Gly	Gly	Val	Ala	Ala	Gly	Ile	Glu	Leu	Ser	
185								190				195				
cag	tgg	ttc	ctg	cta	gtc	gct	gcg	ttt	ggc	tcc	ctg	ttc	atg	gca	tct	739
Gln	Trp	Phe	Leu	Leu	Val	Ala	Ala	Phe	Gly	Ser	Leu	Phe	Met	Ala	Ser	
200								205				210				
gga	aag	cgc	tac	gca	gaa	atc	ctt	ctg	cac	gag	cgc	acc	ggc	gct	aag	787
Gly	Lys	Arg	Tyr	Ala	Glu	Ile	Leu	Leu	His	Glu	Arg	Thr	Gly	Ala	Lys	
215								220				225				
atc	cgc	aag	tcc	ctg	gaa	agc	tac	acc	ccc	acc	tac	ctg	cgc	ttc	gtt	835
Ile	Arg	Lys	Ser	Leu	Glu	Ser	Tyr	Thr	Pro	Thr	Tyr	Leu	Arg	Phe	Val	
230								235				240				245
tgg	acc	atg	gca	gca	aca	gca	gtg	gtc	atg	tcc	tac	gca	ctg	tgg	ggc	883
Trp	Thr	Met	Ala	Ala	Thr	Ala	Val	Val	Met	Ser	Tyr	Ala	Leu	Trp	Gly	
250								255				260				
ttc	gac	ctt	tcc	caa	cac	tcc	acc	gac	gca	ggc	cgc	tgg	tac	caa	atc	931
Phe	Asp	Leu	Ser	Gln	His	Ser	Thr	Asp	Ala	Gly	Pro	Trp	Tyr	Gln	Ile	
265								270				275				
tcc	atg	gtt	cca	ttc	acc	atc	gcc	atc	ctg	cgc	tac	gca	gcc	ggc	gta	979
Ser	Met	Val	Pro	Phe	Thr	Ile	Ala	Ile	Leu	Arg	Tyr	Ala	Ala	Gly	Val	
280								285				290				
gac	acc	ggc	gac	ggc	ggc	ggc	cct	gac	gaa	gtg	gca	ctc	agc	gac	aaa	1027
Asp	Thr	Gly	Asp	Gly	Gly	Ala	Pro	Asp	Glu	Val	Ala	Leu	Ser	Asp	Lys	
295								300				305				
gtt	ctg	cag	gta	cta	gcc	cta	gca	tgg	gtt	ttc	tgc	atc	gtg	atg	gct	1075
Val	Leu	Gln	Val	Leu	Ala	Leu	Ala	Trp	Val	Phe	Cys	Ile	Val	Met	Ala	
310								315				320				325
gtg	tac	atc	atg	cgc	atg	ttt	tgaatatttta	ccaatgaaca	tgc							1119
Val	Tyr	Ile	Met	Pro	Met	Phe										
330																

<210> 28  
 <211> 332  
 <212> PRT  
 <213> *Corynebacterium glutamicum*

<400> 28  
 Val Ser Glu His Ala Ala Glu His His Arg Asp Thr Gln Asn Phe Leu  
 1 5 10 15  
 Thr Ser Glu Pro His Thr Thr Ala Ile Glu Asp Asn Lys Lys Arg Gln  
 20 25 30  
 Pro Pro Lys Asn Leu Ala Asp Gly Met Ile Lys Ala Leu Arg Pro Lys  
 35 40 45  
 Gln Trp Val Lys Asn Val Leu Val Leu Ala Ala Pro Leu Ala Ala Gly  
 50 55 60  
 Ala Asp Ala Ile Phe Asn Gln Arg Thr Ile Ile Asp Val Ala Ile Ala  
 65 70 75 80  
 Phe Val Val Phe Cys Phe Gly Ala Ser Ala Ile Tyr Leu Val Asn Asp  
 85 90 95  
 Ala Arg Asp Val Glu Ala Asp Arg Glu His Pro Thr Lys Arg Phe Arg  
 100 105 110  
 Pro Ile Ala Ala Gly Val Leu Pro Val Gly Met Ala Tyr Gly Met Ala  
 115 120 125  
 Val Ala Leu Ile Ala Leu Ser Ile Gly Leu Ser Phe Leu Ala Thr Asp  
 130 135 140  
 Gly Val Ala Leu Ala Cys Val Ile Gly Val Tyr Ile Ala Leu Gln Leu  
 145 150 155 160  
 Gly Tyr Cys Phe Gly Trp Lys His Met Pro Val Ile Asp Ile Ala Leu  
 165 170 175  
 Val Ser Ser Gly Phe Met Leu Arg Ala Met Ala Gly Gly Val Ala Ala  
 180 185 190  
 Gly Ile Glu Leu Ser Gln Trp Phe Leu Leu Val Ala Ala Phe Gly Ser  
 195 200 205  
 Leu Phe Met Ala Ser Gly Lys Arg Tyr Ala Glu Ile Leu Leu His Glu  
 210 215 220  
 Arg Thr Gly Ala Lys Ile Arg Lys Ser Leu Glu Ser Tyr Thr Pro Thr  
 225 230 235 240  
 Tyr Leu Arg Phe Val Trp Thr Met Ala Ala Thr Ala Val Val Met Ser  
 245 250 255  
 Tyr Ala Leu Trp Gly Phe Asp Leu Ser Gln His Ser Thr Asp Ala Gly  
 260 265 270  
 Pro Trp Tyr Gln Ile Ser Met Val Pro Phe Thr Ile Ala Ile Leu Arg  
 275 280 285



Tyr Ala Ala Gly Val Asp Thr Gly Asp Gly Gly Ala Pro Asp Glu Val  
 290 295 300

Ala Leu Ser Asp Lys Val Leu Gln Val Leu Ala Leu Ala Trp Val Phe  
 305 310 315 320

Cys Ile Val Met Ala Val Tyr Ile Met Pro Met Phe  
 325 330

<210> 29

<211> 2004

<212> DNA

<213> Corynebacterium glutamicum

<220>

<221> CDS

<222> (101)..(1981)

<223> RXA01385

<400> 29

atcacacttt cattcgcaagt gtgatctgaa ctacatttct gggtactgta cggaacacgc 60

tccgtgaatg agataggaaa tcccctcgaa aggaccagac atg cag ttt cat tat 115  
 Met Gln Phe His Tyr  
 1 5

gaa gga tac gca acc ggt gac cca atg gag atg cgc gcg gaa ggt agc 163  
 Glu Gly Tyr Ala Thr Gly Asp Pro Met Glu Met Arg Ala Glu Gly Ser  
 10 15 20

gga atc aac cgc ccg gac gat ctc ccc gag gtc atg gat gtt ctc atc 211  
 Gly Ile Asn Arg Pro Asp Asp Leu Pro Glu Val Met Asp Val Leu Ile  
 25 30 35

gtt ggt gca ggt ccg gct ggc acc atc gca gcg gct cag ctt tcc cga 259  
 Val Gly Ala Gly Pro Ala Gly Thr Ile Ala Ala Ala Gln Leu Ser Arg  
 40 45 50

ttc ccc aat gtg acc acc cgc ctc gta gag aga agc gac cgt cgc ctc 307  
 Phe Pro Asn Val Thr Thr Arg Leu Val Glu Arg Ser Asp Arg Arg Leu  
 55 60 65

gaa cta gcc aat gca gat ggc gtg cac tcc cga acc att gaa act ttc 355  
 Glu Leu Ala Asn Ala Asp Gly Val His Ser Arg Thr Ile Glu Thr Phe  
 70 75 80 85

cag gca ttt ggt ttc gcc cac gag atc ctc gcc gaa gct cat gaa atc 403  
 Gln Ala Phe Gly Phe Ala His Glu Ile Leu Ala Glu Ala His Glu Ile  
 90 95 100

acc gac atg gcg ttc tgg aag ccg gac ccg caa aac cct cgt gag atc 451  
 Thr Asp Met Ala Phe Trp Lys Pro Asp Pro Gln Asn Pro Arg Glu Ile  
 105 110 115

att cgc gac aac agc acc cgc gag ctg cca cag cac atc agt gaa ttt 499  
 Ile Arg Asp Asn Ser Thr Arg Glu Leu Pro Gln His Ile Ser Glu Phe  
 120 125 130

ccg atg gcg ttg ctc acc cag acc cgc atc atc gac cac ttc aac cgg 547

Pro Met Ala Leu Leu Thr	Gln Thr Arg Ile Ile Asp His Phe Asn Arg	
135	140	145
ttc atg aag aac tcc cca acc agg atg aag cct gac tat gga tac gag		595
Phe Met Lys Asn Ser Pro Thr Arg Met Lys Pro Asp Tyr Gly Tyr Glu		
150	155	160
ttc gtg gac ttt gaa gta gaa gaa gac gca gaa tat ccg gta att gtc		643
Phe Val Asp Phe Glu Val Glu Glu Asp Ala Glu Tyr Pro Val Ile Val		
	170	175
acc ctc cgc cgc acc agt ggc gag caa act ggc gaa ttg gtc acc gtc		691
Thr Leu Arg Arg Thr Ser Gly Glu Gln Thr Gly Glu Leu Val Thr Val		
	185	190
cga acc aag tac ctg gtc ggt gcc gat ggt gca cga agc caa gtg cgc		739
Arg Thr Lys Tyr Leu Val Gly Ala Asp Gly Ala Arg Ser Gln Val Arg		
	200	205
aaa tca ctg gga tac cga ctc caa ggt aag cag gct aac cac gct tgg		787
Lys Ser Leu Gly Tyr Arg Leu Gln Gly Lys Gln Ala Asn His Ala Trp		
	215	220
ggt gtg atg gat att cac gcg aac acc gag ttc ccc gac gtg cgc aag		835
Gly Val Met Asp Ile His Ala Asn Thr Glu Phe Pro Asp Val Arg Lys		
	230	235
aag tgc acc atc aaa tct gat tcg ggt cgc acc atc ttg ctc atc cca		883
Lys Cys Thr Ile Lys Ser Asp Ser Gly Arg Thr Ile Leu Leu Ile Pro		
	250	255
cgt gag ggt ggc ttc ctc ttc cgt ctc tac gtt gac ctg ggc gaa gta		931
Arg Glu Gly Gly Phe Leu Phe Arg Leu Tyr Val Asp Leu Gly Glu Val		
	265	270
cct gat gat ggc agc aag gct gtt cgt gat acc cca ctc cag gat gtc		979
Pro Asp Asp Gly Ser Lys Ala Val Arg Asp Thr Pro Leu Gln Asp Val		
	280	285
atc gac acc gcg aac cag atc atg gct cca ttc acc ctc gac gtg aaa		1027
Ile Asp Thr Ala Asn Gln Ile Met Ala Pro Phe Thr Leu Asp Val Lys		
	295	300
aac gtt gtg tgg aac tcc atc tac gag gta ggc cac cgc gtc gca gac		1075
Asn Val Val Trp Asn Ser Ile Tyr Glu Val Gly His Arg Val Ala Asp		
	310	315
cat ttc gat gac cgt gtt tca gaa aaa acc tcg agc gaa cac cca cgc		1123
His Phe Asp Asp Arg Val Ser Glu Lys Thr Ser Ser Glu His Pro Arg		
	330	335
att ttc att gct ggc gac gcc tgc cac acc cac agc gct aag gct ggc		1171
Ile Phe Ile Ala Gly Asp Ala Cys His Thr His Ser Ala Lys Ala Gly		
	345	350
cag ggc atg aac gtg tcc atg cag gac gga ttc aac ctt ggc tgg aag		1219
Gln Gly Met Asn Val Ser Met Gln Asp Gly Phe Asn Leu Gly Trp Lys		
	360	365
ctt gga cat gtg gcc agc gga aat agc cca cgc gaa cta ctt cag acc		1267
Leu Gly His Val Ala Ser Gly Asn Ser Pro Arg Glu Leu Leu Gln Thr		

375	380	385	
tac gct gaa gag cgc gaa gac att gcc tac aag ctc atc gag tac gac Tyr Ala Glu Glu Arg Glu Asp Ile Ala Tyr Lys Leu Ile Glu Tyr Asp 390 395 400 405			1315
aag aac tgg tca aca ctc atg gca aag cca agc agc gaa atg ggc agt Lys Asn Trp Ser Thr Leu Met Ala Lys Pro Ser Ser Glu Met Gly Ser 410 415 420			1363
gcc caa gac ctt gag gat ttc tac cgc gcg aac tct gag ttc aat gcc Ala Gln Asp Leu Glu Asp Phe Tyr Arg Ala Asn Ser Glu Phe Asn Ala 425 430 435			1411
ggc tac atg acc cac tat cct cct tct tcc atc aca atg gat ggc agc Gly Tyr Met Thr His Tyr Pro Pro Ser Ser Ile Thr Met Asp Gly Ser 440 445 450			1459
aac caa gat ctg gca aag ggc tac cca att ggc cga cgc ttc aag tca Asn Gln Asp Leu Ala Lys Gly Tyr Pro Ile Gly Arg Arg Phe Lys Ser 455 460 465			1507
gcg atg gtt ggt cga gtc tgc gac ttc acc gaa aca cac ctc ggt cac Ala Met Val Gly Arg Val Cys Asp Phe Thr Glu Thr His Leu Gly His 470 475 480 485			1555
caa gca aca gcc gac gga cgc atg cgc gca tac gtc ttc gca gga tcc Gln Ala Thr Ala Asp Gly Arg Met Arg Ala Tyr Val Phe Ala Gly Ser 490 495 500			1603
gat gca ctt aac ggc gag ggt tct gag cta gac cgc tgg gca gaa tgg Asp Ala Leu Asn Gly Glu Gly Ser Glu Leu Asp Arg Trp Ala Glu Trp 505 510 515			1651
gca gag gcg aac ctt gac ccc acg ctt gtc gac gcc aag gtg att tac Ala Glu Ala Asn Leu Asp Pro Thr Leu Val Asp Ala Lys Val Ile Tyr 520 525 530			1699
caa agc cct tat acc gag ctc gac acc cgc cag gtt cca tcc gtg ttc Gln Ser Pro Tyr Thr Glu Leu Asp Thr Arg Gln Val Pro Ser Val Phe 535 540 545			1747
aaa cct gca gtc ggg atc ttc gaa ctg acc aat gtg gaa aac tcc ttc Lys Pro Ala Val Gly Ile Phe Glu Leu Thr Asn Val Glu Asn Ser Phe 550 555 560 565			1795
ggt atc acc acg gac tcc gac atc ttt gat agt cgc gag atc tcc cgc Gly Ile Thr Thr Asp Ser Asp Ile Phe Asp Ser Arg Glu Ile Ser Arg 570 575 580			1843
gat ggt gtc gtg gtg gta gtc cga cca gac caa tac gtt tcc gga atc Asp Gly Val Val Val Val Arg Pro Asp Gln Tyr Val Ser Gly Ile 585 590 595			1891
ttc cca ctc act gat acc caa ggg ctt ggc gaa ttc ctc acc gga tac Phe Pro Leu Thr Asp Thr Gln Gly Leu Gly Glu Phe Leu Thr Gly Tyr 600 605 610			1939
ttc ccc aaa atg aaa ggc gca cat cag cta atc aac gcg aac Phe Pro Lys Met Lys Gly Ala His Gln Leu Ile Asn Ala Asn 615 620 625			1981

taaggcacag ctgttaaaac agt

2004

&lt;210&gt; 30

&lt;211&gt; 627

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 30

Met	Gln	Phe	His	Tyr	Glu	Gly	Tyr	Ala	Thr	Gly	Asp	Pro	Met	Glu	Met
1				5					10					15	

Arg	Ala	Glu	Gly	Ser	Gly	Ile	Asn	Arg	Pro	Asp	Asp	Leu	Pro	Glu	Val
		20					25						30		

Met	Asp	Val	Leu	Ile	Val	Gly	Ala	Gly	Pro	Ala	Gly	Thr	Ile	Ala	Ala
		35				40						45			

Ala	Gln	Leu	Ser	Arg	Phe	Pro	Asn	Val	Thr	Thr	Arg	Leu	Val	Glu	Arg
	50					55					60				

Ser	Asp	Arg	Arg	Leu	Glu	Leu	Ala	Asn	Ala	Asp	Gly	Val	His	Ser	Arg
65				70						75					80

Thr	Ile	Glu	Thr	Phe	Gln	Ala	Phe	Gly	Phe	Ala	His	Glu	Ile	Leu	Ala
				85				90						95	

Glu	Ala	His	Glu	Ile	Thr	Asp	Met	Ala	Phe	Trp	Lys	Pro	Asp	Pro	Gln
		100						105					110		

Asn	Pro	Arg	Glu	Ile	Ile	Arg	Asp	Asn	Ser	Thr	Arg	Glu	Leu	Pro	Gln
		115				120						125			

His	Ile	Ser	Glu	Phe	Pro	Met	Ala	Leu	Leu	Thr	Gln	Thr	Arg	Ile	Ile
	130					135					140				

Asp	His	Phe	Asn	Arg	Phe	Met	Lys	Asn	Ser	Pro	Thr	Arg	Met	Lys	Pro
145					150					155					160

Asp	Tyr	Gly	Tyr	Glu	Phe	Val	Asp	Phe	Glu	Val	Glu	Glu	Asp	Ala	Glu
				165					170					175	

Tyr	Pro	Val	Ile	Val	Thr	Leu	Arg	Arg	Thr	Ser	Gly	Glu	Gln	Thr	Gly
		180						185					190		

Glu	Leu	Val	Thr	Val	Arg	Thr	Lys	Tyr	Leu	Val	Gly	Ala	Asp	Gly	Ala
		195					200					205			

Arg	Ser	Gln	Val	Arg	Lys	Ser	Leu	Gly	Tyr	Arg	Leu	Gln	Gly	Lys	Gln
	210					215					220				

Ala	Asn	His	Ala	Trp	Gly	Val	Met	Asp	Ile	His	Ala	Asn	Thr	Glu	Phe
225					230					235					240

Pro	Asp	Val	Arg	Lys	Lys	Cys	Thr	Ile	Lys	Ser	Asp	Ser	Gly	Arg	Thr
				245					250					255	

Ile	Leu	Leu	Ile	Pro	Arg	Glu	Gly	Gly	Phe	Leu	Phe	Arg	Leu	Tyr	Val
			260					265						270	

Asp Leu Gly Glu Val Pro Asp Asp Gly Ser Lys Ala Val Arg Asp Thr  
 275 280 285  
 Pro Leu Gln Asp Val Ile Asp Thr Ala Asn Gln Ile Met Ala Pro Phe  
 290 295 300  
 Thr Leu Asp Val Lys Asn Val Val Trp Asn Ser Ile Tyr Glu Val Gly  
 305 310 315 320  
 His Arg Val Ala Asp His Phe Asp Asp Arg Val Ser Glu Lys Thr Ser  
 325 330 335  
 Ser Glu His Pro Arg Ile Phe Ile Ala Gly Asp Ala Cys His Thr His  
 340 345 350  
 Ser Ala Lys Ala Gly Gln Gly Met Asn Val Ser Met Gln Asp Gly Phe  
 355 360 365  
 Asn Leu Gly Trp Lys Leu Gly His Val Ala Ser Gly Asn Ser Pro Arg  
 370 375 380  
 Glu Leu Leu Gln Thr Tyr Ala Glu Glu Arg Glu Asp Ile Ala Tyr Lys  
 385 390 395 400  
 Leu Ile Glu Tyr Asp Lys Asn Trp Ser Thr Leu Met Ala Lys Pro Ser  
 405 410 415  
 Ser Glu Met Gly Ser Ala Gln Asp Leu Glu Asp Phe Tyr Arg Ala Asn  
 420 425 430  
 Ser Glu Phe Asn Ala Gly Tyr Met Thr His Tyr Pro Pro Ser Ser Ile  
 435 440 445  
 Thr Met Asp Gly Ser Asn Gln Asp Leu Ala Lys Gly Tyr Pro Ile Gly  
 450 455 460  
 Arg Arg Phe Lys Ser Ala Met Val Gly Arg Val Cys Asp Phe Thr Glu  
 465 470 475 480  
 Thr His Leu Gly His Gln Ala Thr Ala Asp Gly Arg Met Arg Ala Tyr  
 485 490 495  
 Val Phe Ala Gly Ser Asp Ala Leu Asn Gly Glu Gly Ser Glu Leu Asp  
 500 505 510  
 Arg Trp Ala Glu Trp Ala Glu Ala Asn Leu Asp Pro Thr Leu Val Asp  
 515 520 525  
 Ala Lys Val Ile Tyr Gln Ser Pro Tyr Thr Glu Leu Asp Thr Arg Gln  
 530 535 540  
 Val Pro Ser Val Phe Lys Pro Ala Val Gly Ile Phe Glu Leu Thr Asn  
 545 550 555 560  
 Val Glu Asn Ser Phe Gly Ile Thr Thr Asp Ser Asp Ile Phe Asp Ser  
 565 570 575  
 Arg Glu Ile Ser Arg Asp Gly Val Val Val Val Arg Pro Asp Gln  
 580 585 590  
 Tyr Val Ser Gly Ile Phe Pro Leu Thr Asp Thr Gln Gly Leu Gly Glu

38

aac cgt ttg acc gat gtc tcc cac gct ctc gag gtc gca acc cgc aag 595  
 Asn Arg Leu Thr Asp Val Ser His Ala Leu Glu Val Ala Thr Arg Lys  
 150 155 160 165  
 gct gag tcc aag ttc ggc gtc gcg ctc ggc atc gtc gat ggc tac ggc 643  
 Ala Glu Ser Lys Phe Gly Val Ala Leu Gly Ile Val Asp Gly Tyr Gly  
 170 175 180  
 gga cac ggc att ggc cgc cac atg cac gag gag cca tac ttg gct aat 691  
 Gly His Gly Ile Gly Arg His Met His Glu Glu Pro Tyr Leu Ala Asn  
 185 190 195  
 gag ggc aag gcc ggc aag ggc cct gtg att cag gag ggc tcc gtg ctc 739  
 Glu Gly Lys Ala Gly Lys Gly Pro Val Ile Gln Glu Gly Ser Val Leu  
 200 205 210  
 gcc att gag cct atg ctc acc ctc ggc acc gaa gat tcc gca gtg ctg 787  
 Ala Ile Glu Pro Met Leu Thr Leu Gly Thr Glu Asp Ser Ala Val Leu  
 215 220 225  
 gaa gat gat tgg act gtc gtg act ctc gac ggt tca tgg gca tca cac 835  
 Glu Asp Asp Trp Thr Val Val Thr Leu Asp Gly Ser Trp Ala Ser His  
 230 235 240 245  
 tgg gag cac acc gtt gca gcc acc aag ggc ggc ccg cgc atc ctc acg 883  
 Trp Glu His Thr Val Ala Ala Thr Lys Gly Gly Pro Arg Ile Leu Thr  
 250 255 260  
 ccg cgt tat taaaatgatg cttttcgacg cat 915  
 Pro Arg Tyr

&lt;210&gt; 32

&lt;211&gt; 264

&lt;212&gt; PRT

<213> *Corynebacterium glutamicum*

&lt;400&gt; 32

Met Gly Phe Arg Ser Lys Lys Lys Val Ile Ala Ala Lys Thr Ala Ala  
 1 5 10 15  
 Glu Leu Asp Ala Met Gln Ala Ala Gly Glu Ile Val Gly Lys Ala Leu  
 20 25 30  
 Gln Ala Val Arg Ala Glu Ala Lys Ala Gly Met Ser Thr Trp Asp Leu  
 35 40 45  
 Asp Gln Ile Ala Glu Gln Val Ile Arg Asp Ala Gly Ala Val Pro Thr  
 50 55 60  
 Phe Leu Gly Tyr Gln Gly Phe Pro Ala Ser Val Cys Ala Ser Val Asn  
 65 70 75 80  
 Glu Val Ile Val His Gly Ile Pro Ser Lys Glu Thr Ile Leu Glu Glu  
 85 90 95  
 Gly Asp Leu Val Ser Ile Asp Cys Gly Ala Thr Phe Asp Gly Trp Val  
 100 105 110

Gly Asp Ser Ala Trp Ser Phe Gly Ile Gly Glu Leu Asp Glu Asp Val  
 115 120 125  
 Gln Gly Leu Asn Leu Ala Thr Glu Trp Val Leu Met Glu Gly Met Lys  
 130 135 140  
 Ala Met Val Pro Gly Asn Arg Leu Thr Asp Val Ser His Ala Leu Glu  
 145 150 155 160  
 Val Ala Thr Arg Lys Ala Glu Ser Lys Phe Gly Val Ala Leu Gly Ile  
 165 170 175  
 Val Asp Gly Tyr Gly Gly His Gly Ile Gly Arg His Met His Glu Glu  
 180 185 190  
 Pro Tyr Leu Ala Asn Glu Gly Lys Ala Gly Lys Gly Pro Val Ile Gln  
 195 200 205  
 Glu Gly Ser Val Leu Ala Ile Glu Pro Met Leu Thr Leu Gly Thr Glu  
 210 215 220  
 Asp Ser Ala Val Leu Glu Asp Asp Trp Thr Val Val Thr Leu Asp Gly  
 225 230 235 240  
 Ser Trp Ala Ser His Trp Glu His Thr Val Ala Ala Thr Lys Gly Gly  
 245 250 255  
 Pro Arg Ile Leu Thr Pro Arg Tyr  
 260

<210> 33  
 <211> 506  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (1)..(483)  
 <223> FRXA00675

<400> 33  
 tgc ggc gca acc ttt gat ggt tgg gtc ggc gat tcc gcg tgg agc ttc 48  
 Cys Gly Ala Thr Phe Asp Gly Trp Val Gly Asp Ser Ala Trp Ser Phe  
 1 5 10 15  
 ggc atc ggc gag ctg gac gag gac gtc cag ggt ctc aac ttg gct acc 96  
 Gly Ile Gly Glu Leu Asp Glu Asp Val Gln Gly Leu Asn Leu Ala Thr  
 20 25 30  
 gag tgg gtc ctc atg gaa ggc atg aag gcc atg gtt cca ggc aac cgt 144  
 Glu Trp Val Leu Met Glu Gly Met Lys Ala Met Val Pro Gly Asn Arg  
 35 40 45  
 ttg acc gat gtc tcc cac gct ctc gag gtc gca acc cgc aag gct gag 192  
 Leu Thr Asp Val Ser His Ala Leu Glu Val Ala Thr Arg Lys Ala Glu  
 50 55 60  
 tcc aag ttc ggc gtc gcg ctc ggc atc gtc gat ggc tac ggc gga cac 240  
 Ser Lys Phe Gly Val Ala Leu Gly Ile Val Asp Gly Tyr Gly Gly His  
 65 70 75 80



ggc att ggc cgc cac atg cac gag gag cca tac ttg gct aat gag ggc 288  
 Gly Ile Gly Arg His Met His Glu Glu Pro Tyr Leu Ala Asn Glu Gly  
                     85                    90                    95  
  
 aag gcc ggc aag ggc cct gtg att cag gag ggc tcc gtg ctc gcc att 336  
 Lys Ala Gly Lys Gly Pro Val Ile Gln Glu Gly Ser Val Leu Ala Ile  
                     100                    105                    110  
  
 gag cct atg ctc acc ctc ggc acc gaa gat tcc gca gtg ctg gaa gat 384  
 Glu Pro Met Leu Thr Leu Gly Thr Glu Asp Ser Ala Val Leu Glu Asp  
                     115                    120                    125  
  
 gat tgg act gtc gtg act ctc gac ggt tca tgg gca tca cac tgg gag 432  
 Asp Trp Thr Val Val Thr Leu Asp Gly Ser Trp Ala Ser His Trp Glu  
                     130                    135                    140  
  
 cac acc gtt gca gcc acc aag ggc ggc ccg cgc atc ctc acg ccg cgt 480  
 His Thr Val Ala Ala Thr Lys Gly Gly Pro Arg Ile Leu Thr Pro Arg  
 145                    150                    155                    160  
  
 tat taaaatgatg cttttcgacg cat 506  
 Tyr

<210> 34  
 <211> 161  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 34  
 Cys Gly Ala Thr Phe Asp Gly Trp Val Gly Asp Ser Ala Trp Ser Phe  
   1                    5                    10                    15  
  
 Gly Ile Gly Glu Leu Asp Glu Asp Val Gln Gly Leu Asn Leu Ala Thr  
                     20                    25                    30  
  
 Glu Trp Val Leu Met Glu Gly Met Lys Ala Met Val Pro Gly Asn Arg  
                     35                    40                    45  
  
 Leu Thr Asp Val Ser His Ala Leu Glu Val Ala Thr Arg Lys Ala Glu  
                     50                    55                    60  
  
 Ser Lys Phe Gly Val Ala Leu Gly Ile Val Asp Gly Tyr Gly Gly His  
                     65                    70                    75                    80  
  
 Gly Ile Gly Arg His Met His Glu Glu Pro Tyr Leu Ala Asn Glu Gly  
                     85                    90                    95  
  
 Lys Ala Gly Lys Gly Pro Val Ile Gln Glu Gly Ser Val Leu Ala Ile  
                     100                    105                    110  
  
 Glu Pro Met Leu Thr Leu Gly Thr Glu Asp Ser Ala Val Leu Glu Asp  
                     115                    120                    125  
  
 Asp Trp Thr Val Val Thr Leu Asp Gly Ser Trp Ala Ser His Trp Glu  
                     130                    135                    140  
  
 His Thr Val Ala Ala Thr Lys Gly Gly Pro Arg Ile Leu Thr Pro Arg  
 145                    150                    155                    160



gaa gaa gcc atg atg cgt tcc atc cgt gca gca aag cct gga cgt gaa 643  
 Glu Glu Ala Met Met Arg Ser Ile Arg Ala Ala Lys Pro Gly Arg Glu  
 170 175 180  
 atc aac gtc att ggg cgt gtc att gag tct tac gcc aag cgt ttt ggc 691  
 Ile Asn Val Ile Gly Arg Val Ile Glu Ser Tyr Ala Lys Arg Phe Gly  
 185 190 195  
 tac aac gtg gtc cgc gat ttc acc gga cac ggc atc ggc cca act ttc 739  
 Tyr Asn Val Val Arg Asp Phe Thr Gly His Gly Ile Gly Pro Thr Phe  
 200 205 210  
 cac aac ggc ctt gtg gtg ctg cac tac gac aac act cag tac cgc gat 787  
 His Asn Gly Leu Val Val Leu His Tyr Asp Asn Thr Gln Tyr Arg Asp  
 215 220 225  
 ctg ctc gtg cca ggc atg acc ttg acc atc gag cca atg atc aac ctt 835  
 Leu Leu Val Pro Gly Met Thr Leu Thr Ile Glu Pro Met Ile Asn Leu  
 230 235 240 245  
 ggt tcc ctc gac tac gag atc tgg gaa gat gat tgg act gtc caa aac 883  
 Gly Ser Leu Asp Tyr Glu Ile Trp Glu Asp Asp Trp Thr Val Gln Asn  
 250 255 260  
 gtt gac cgt aag ttc agc gcg cag ttc gag cac acc att gtc atc acc 931  
 Val Asp Arg Lys Phe Ser Ala Gln Phe Glu His Thr Ile Val Ile Thr  
 265 270 275  
 gaa gac ggc aat gag atc ctc acc ctc cca gac gat tcc gtc 973  
 Glu Asp Gly Asn Glu Ile Leu Thr Leu Pro Asp Asp Ser Val  
 280 285 290  
 taaaaacgcc taggccacaa gcc 996

<210> 36  
 <211> 291  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 36  
 Met Ser Lys Met Arg Ala Pro Leu Val Pro Gly Ile Pro Thr Pro Ile  
 1 5 10 15  
 Arg Glu Val Pro Ala His Ile Glu Arg Pro Glu Tyr Val Trp Lys Asp  
 20 25 30  
 Glu Val Gln Glu Ala Ile Gly Glu Pro Phe Val Gln Ala Pro Glu Val  
 35 40 45  
 Ile Glu Lys Met Arg Glu Thr Ser Arg Ile Ala Ala Asn Ser Leu Lys  
 50 55 60  
 Ile Ala Gly Glu Ala Val Lys Pro Gly Val Thr Thr Asp Glu Leu Asp  
 65 70 75 80  
 Arg Ile Val His Glu Tyr Thr Cys Asp Met Gly Ala Tyr Pro Ser Asp  
 85 90 95  
 Leu Gly Tyr Arg Gly Phe Thr Lys Ser Ser Cys Ile Ser Leu Asn Glu

100	105	110
Ile Val Cys His Gly Ile Pro Asp Ser Thr Val Ile Glu Glu Gly Asp 115 120 125		
Ile Val Asn Ile Asp Val Thr Ala Phe Lys His Gly Val His Gly Asp 130 135 140		
Cys Asn Ala Thr Phe Leu Ala Gly Asp Val Ser Glu Glu His Arg Leu 145 150 155 160		
Leu Val Glu Arg Thr Glu Glu Ala Met Met Arg Ser Ile Arg Ala Ala 165 170 175		
Lys Pro Gly Arg Glu Ile Asn Val Ile Gly Arg Val Ile Glu Ser Tyr 180 185 190		
Ala Lys Arg Phe Gly Tyr Asn Val Val Arg Asp Phe Thr Gly His Gly 195 200 205		
Ile Gly Pro Thr Phe His Asn Gly Leu Val Val Leu His Tyr Asp Asn 210 215 220		
Thr Gln Tyr Arg Asp Leu Leu Val Pro Gly Met Thr Leu Thr Ile Glu 225 230 235 240		
Pro Met Ile Asn Leu Gly Ser Leu Asp Tyr Glu Ile Trp Glu Asp Asp 245 250 255		
Trp Thr Val Gln Asn Val Asp Arg Lys Phe Ser Ala Gln Phe Glu His 260 265 270		
Thr Ile Val Ile Thr Glu Asp Gly Asn Glu Ile Leu Thr Leu Pro Asp 275 280 285		
Asp Ser Val 290		

<210> 37  
 <211> 1644  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1621)  
 <223> RXA01358

<400> 37  
 gagggtaaac cttataactcc gattcggtttt gatgtggtgg caatcggtttt ggatccgcac 60  
 acgggggcgcc cagagatcac tgtttatgag gatgtggaac atg gcg ctc ggt aga 115  
 Met Ala Leu Gly Arg  
 1 5  
 act ata tcc act gcg caa tta ggt gtg cag gca aaa atc gtt cgt gtg 163  
 Thr Ile Ser Thr Ala Gln Leu Gly Val Gln Ala Lys Ile Val Arg Val  
 10 15 20  
 gag gct aat gtt ggc cca gga ttg cct ggt acc tac att gtt gga tta 211

Glu	Ala	Asn	Val	Gly	Pro	Gly	Leu	Pro	Gly	Thr	Tyr	Ile	Val	Gly	Leu		
			25					30					35				
gca	gat	act	gcc	att	tca	gaa	tct	cgg	gat	cgt	att	aaa	act	gcg	gtg	259	
Ala	Asp	Thr	Ala	Ile	Ser	Glu	Ser	Arg	Asp	Arg	Ile	Lys	Thr	Ala	Val		
		40					45					50					
caa	aac	agt	ggc	ttg	atg	tgg	cca	aag	acc	aaa	gtg	atc	att	aac	ctc	307	
Gln	Asn	Ser	Gly	Leu	Met	Trp	Pro	Lys	Thr	Lys	Val	Ile	Ile	Asn	Leu		
		55				60					65						
tcg	ccg	gct	tcc	atg	cgt	aaa	caa	ggc	tcg	cag	tgt	gat	ctc	gcg	atg	355	
Ser	Pro	Ala	Ser	Met	Arg	Lys	Gln	Gly	Ser	Gln	Cys	Asp	Leu	Ala	Met		
70					75					80					85		
acc	gtt	gca	gtt	ctc	gtt	gcc	cat	ggc	tct	aac	ccc	aaa	gcg	aag	ttt	403	
Thr	Val	Ala	Val	Leu	Val	Ala	His	Gly	Ser	Asn	Pro	Lys	Ala	Lys	Phe		
				90					95					100			
cat	gcg	cag	aac	acg	tta	ttt	ctg	ggc	gag	gtg	gcg	ctt	gat	gga	acc	451	
His	Ala	Gln	Asn	Thr	Leu	Phe	Leu	Gly	Glu	Val	Ala	Leu	Asp	Gly	Thr		
			105					110					115				
ttg	ctc	cct	gtt	act	gga	gtt	ctt	cca	gcg	ctg	ttg	gcc	gcg	aag	gag	499	
Leu	Leu	Pro	Val	Thr	Gly	Val	Leu	Pro	Ala	Leu	Leu	Ala	Ala	Lys	Glu		
		120					125					130					
gaa	ggc	att	ggc	aag	att	gtg	atc	ccg	gag	gga	aat	gcc	caa	gaa	gca	547	
Glu	Gly	Ile	Gly	Lys	Ile	Val	Ile	Pro	Glu	Gly	Asn	Ala	Gln	Glu	Ala		
		135				140					145						
gga	ctg	gtt	gag	gat	cct	tca	gtc	ttt	ttg	gca	cat	tcc	atc	gac	cag	595	
Gly	Leu	Val	Glu	Asp	Pro	Ser	Val	Phe	Leu	Ala	His	Ser	Ile	Asp	Gln		
150					155					160					165		
gtc	ttg	cga	tgg	ctt	gat	ggg	gaa	gag	gcg	ttg	cct	cag	cct	gga	cta	643	
Val	Leu	Arg	Trp	Leu	Asp	Gly	Glu	Glu	Ala	Leu	Pro	Gln	Pro	Gly	Leu		
				170					175					180			
ttc	aat	gat	gaa	aat	agc	ctc	aaa	ctg	cct	gat	atg	cgt	gat	gtg	gtg	691	
Phe	Asn	Asp	Glu	Asn	Ser	Leu	Lys	Leu	Pro	Asp	Met	Arg	Asp	Val	Val		
			185					190					195				
gga	caa	cca	gaa	gct	agg	ttt	gct	gca	gaa	gta	gct	gct	gcc	ggc	ggc	739	
Gly	Gln	Pro	Glu	Ala	Arg	Phe	Ala	Ala	Glu	Val	Ala	Ala	Ala	Gly	Gly		
		200					205					210					
cac	cac	atg	ctg	atg	att	ggc	cct	ccc	ggc	tct	gga	aaa	tcc	atg	atc	787	
His	His	Met	Leu	Met	Ile	Gly	Pro	Pro	Gly	Ser	Gly	Lys	Ser	Met	Ile		
		215				220					225						
gcc	gag	cgg	att	ccc	agc	ttg	ctt	ccc	gaa	ctg	tcg	ccc	caa	caa	atg	835	
Ala	Glu	Arg	Ile	Pro	Ser	Leu	Leu	Pro	Glu	Leu	Ser	Pro	Gln	Gln	Met		
230					235					240					245		
atc	gag	gcg	acg	gca	gtg	cat	tcc	gtt	gtg	ggg	cga	acc	ttt	tca	ggg	883	
Ile	Glu	Ala	Thr	Ala	Val	His	Ser	Val	Val	Gly	Arg	Thr	Phe	Ser	Gly		
				250					255					260			
ccg	gtg	tcg	agg	gct	ccg	ttt	att	tcc	cca	cac	cac	aat	gtc	agc	aag	931	
Pro	Val	Ser	Arg	Ala	Pro	Phe	Ile	Ser	Pro	His	His	Asn	Val	Ser	Lys		

265	270	275	
gct gct ctg ctc gga ggt ggc tcg gga tct cct ctt ccg ggt gct atc Ala Ala Leu Leu Gly Gly Gly Ser Gly Ser Pro Leu Pro Gly Ala Ile 280 285 290			979
agc ctg gcg cat cat ggt gtg ttg ttc ctg gat gag gtt agt gag att Ser Leu Ala His His Gly Val Leu Phe Leu Asp Glu Val Ser Glu Ile 295 300 305			1027
cca gcg tca atc ctt gat tct ttg agg act cca ttg gaa tac ggc tcg Pro Ala Ser Ile Leu Asp Ser Leu Arg Thr Pro Leu Glu Tyr Gly Ser 310 315 320 325			1075
atc cgc atc atc aga tcc cgc cat gat gtc acc ttc ccc gca cag ttc Ile Arg Ile Ile Arg Ser Arg His Asp Val Thr Phe Pro Ala Gln Phe 330 335 340			1123
cag ctc atc ctc gcg gcc aat ccg tgt aga tgc ggt gca gaa cag cct Gln Leu Ile Leu Ala Ala Asn Pro Cys Arg Cys Gly Ala Glu Gln Pro 345 350 355			1171
caa gaa tgt gtc tgt tct ggc tca gct cgc gcg acg tac ctt aat aat Gln Glu Cys Val Cys Ser Gly Ser Ala Arg Ala Thr Tyr Leu Asn Asn 360 365 370			1219
ctt tcg ggt ccg ttg agg gat cgc ttg gac atg gtt gtt gcc acc cac Leu Ser Gly Pro Leu Arg Asp Arg Leu Asp Met Val Val Ala Thr His 375 380 385			1267
tct aaa ggt gca gtg ctg cgt agt gat gac gtt gag gca tct gct ccc Ser Lys Gly Ala Val Leu Arg Ser Asp Asp Val Glu Ala Ser Ala Pro 390 395 400 405			1315
att gct gat cgg gtg gca caa gct cgt gag agg gca gct ttc cga tgg Ile Ala Asp Arg Val Ala Gln Ala Arg Glu Arg Ala Ala Phe Arg Trp 410 415 420			1363
cgc cgt tct gga ctg gga aat ctt gtt aat gca cac gta gat cca cac Arg Arg Ser Gly Leu Gly Asn Leu Val Asn Ala His Val Asp Pro His 425 430 435			1411
ttc ttg cgg agg aat ttt gcc gca aca gaa gac gcc atg gtg tac ctc Phe Leu Arg Arg Asn Phe Ala Ala Thr Glu Asp Ala Met Val Tyr Leu 440 445 450			1459
ggc gcg ttc ctg gcg gaa gga aca atc tcc caa cgt ggc tgc gat cgg Gly Ala Phe Leu Ala Glu Gly Thr Ile Ser Gln Arg Gly Cys Asp Arg 455 460 465			1507
gcc ata aaa ctg ggt tgg acc ttg tgc gat ttg gat ggg gaa cag cag Ala Ile Lys Leu Gly Trp Thr Leu Cys Asp Leu Asp Gly Glu Gln Gln 470 475 480 485			1555
ccc aat ctt gac cat att gcg cga gcc atg gag ctt cgg ggc act aca Pro Asn Leu Asp His Ile Ala Arg Ala Met Glu Leu Arg Gly Thr Thr 490 495 500			1603
tac agt gag gta gca gca tgattgactc ccgcttgctg gca Tyr Ser Glu Val Ala Ala 505			1644

<210> 38  
 <211> 507  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 38

```

Met Ala Leu Gly Arg Thr Ile Ser Thr Ala Gln Leu Gly Val Gln Ala
 1           5           10           15

Lys Ile Val Arg Val Glu Ala Asn Val Gly Pro Gly Leu Pro Gly Thr
          20           25           30

Tyr Ile Val Gly Leu Ala Asp Thr Ala Ile Ser Glu Ser Arg Asp Arg
          35           40           45

Ile Lys Thr Ala Val Gln Asn Ser Gly Leu Met Trp Pro Lys Thr Lys
          50           55           60

Val Ile Ile Asn Leu Ser Pro Ala Ser Met Arg Lys Gln Gly Ser Gln
          65           70           75           80

Cys Asp Leu Ala Met Thr Val Ala Val Leu Val Ala His Gly Ser Asn
          85           90           95

Pro Lys Ala Lys Phe His Ala Gln Asn Thr Leu Phe Leu Gly Glu Val
          100           105           110

Ala Leu Asp Gly Thr Leu Leu Pro Val Thr Gly Val Leu Pro Ala Leu
          115           120           125

Leu Ala Ala Lys Glu Glu Gly Ile Gly Lys Ile Val Ile Pro Glu Gly
          130           135           140

Asn Ala Gln Glu Ala Gly Leu Val Glu Asp Pro Ser Val Phe Leu Ala
          145           150           155           160

His Ser Ile Asp Gln Val Leu Arg Trp Leu Asp Gly Glu Glu Ala Leu
          165           170           175

Pro Gln Pro Gly Leu Phe Asn Asp Glu Asn Ser Leu Lys Leu Pro Asp
          180           185           190

Met Arg Asp Val Val Gly Gln Pro Glu Ala Arg Phe Ala Ala Glu Val
          195           200           205

Ala Ala Ala Gly Gly His His Met Leu Met Ile Gly Pro Pro Gly Ser
          210           215           220

Gly Lys Ser Met Ile Ala Glu Arg Ile Pro Ser Leu Leu Pro Glu Leu
          225           230           235           240

Ser Pro Gln Gln Met Ile Glu Ala Thr Ala Val His Ser Val Val Gly
          245           250           255

Arg Thr Phe Ser Gly Pro Val Ser Arg Ala Pro Phe Ile Ser Pro His
          260           265           270

His Asn Val Ser Lys Ala Ala Leu Leu Gly Gly Gly Ser Gly Ser Pro
          275           280           285

```

Leu Pro Gly Ala Ile Ser Leu Ala His His Gly Val Leu Phe Leu Asp  
 290 295 300  
 Glu Val Ser Glu Ile Pro Ala Ser Ile Leu Asp Ser Leu Arg Thr Pro  
 305 310 315 320  
 Leu Glu Tyr Gly Ser Ile Arg Ile Ile Arg Ser Arg His Asp Val Thr  
 325 330 335  
 Phe Pro Ala Gln Phe Gln Leu Ile Leu Ala Ala Asn Pro Cys Arg Cys  
 340 345 350  
 Gly Ala Glu Gln Pro Gln Glu Cys Val Cys Ser Gly Ser Ala Arg Ala  
 355 360 365  
 Thr Tyr Leu Asn Asn Leu Ser Gly Pro Leu Arg Asp Arg Leu Asp Met  
 370 375 380  
 Val Val Ala Thr His Ser Lys Gly Ala Val Leu Arg Ser Asp Asp Val  
 385 390 395 400  
 Glu Ala Ser Ala Pro Ile Ala Asp Arg Val Ala Gln Ala Arg Glu Arg  
 405 410 415  
 Ala Ala Phe Arg Trp Arg Arg Ser Gly Leu Gly Asn Leu Val Asn Ala  
 420 425 430  
 His Val Asp Pro His Phe Leu Arg Arg Asn Phe Ala Ala Thr Glu Asp  
 435 440 445  
 Ala Met Val Tyr Leu Gly Ala Phe Leu Ala Glu Gly Thr Ile Ser Gln  
 450 455 460  
 Arg Gly Cys Asp Arg Ala Ile Lys Leu Gly Trp Thr Leu Cys Asp Leu  
 465 470 475 480  
 Asp Gly Glu Gln Gln Pro Asn Leu Asp His Ile Ala Arg Ala Met Glu  
 485 490 495  
 Leu Arg Gly Thr Thr Tyr Ser Glu Val Ala Ala  
 500 505

<210> 39  
 <211> 1173  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1150)  
 <223> RXA01458

<400> 39  
 gccattagag ttcattattta ccaatctaca aggcaactat ttcctattgc tacgatccga 60  
 acaacgctgt aagcgaacac aggttgtagg gtggatattc gtg aat cgc cga atc 115  
 Val Asn Arg Arg Ile  
 1 5



aag act ctg acg tgg ggt gct atc cct ttg gtg ctg ctg gca tcg ttg	163
Lys Thr Leu Thr Trp Gly Ala Ile Pro Leu Val Leu Leu Ala Ser Leu	
10 15 20	
gta agc att gac cat att ccg gga aca aac atc aac ttg agc gtg cct	211
Val Ser Ile Asp His Ile Pro Gly Thr Asn Ile Asn Leu Ser Val Pro	
25 30 35	
tat gcc gct gaa ggc cca ggt cct acg atc aat acg ctt ggt cag gtc	259
Tyr Ala Ala Glu Gly Pro Gly Pro Thr Ile Asn Thr Leu Gly Gln Val	
40 45 50	
gac ggc gag gat gtt gtg tcc atc agt agt gct gat ctg gat gag acc	307
Asp Gly Glu Asp Val Val Ser Ile Ser Ser Ala Asp Leu Asp Glu Thr	
55 60 65	
gaa ggt aac ctg aac atg acc act gtg tcg gtt cgt tcc ggc atg aca	355
Glu Gly Asn Leu Asn Met Thr Thr Val Ser Val Arg Ser Gly Met Thr	
70 75 80 85	
ttg tcg cag gta att tcc cga tgg ctg ttt acc gat gac aca atc gtt	403
Leu Ser Gln Val Ile Ser Arg Trp Leu Phe Thr Asp Asp Thr Ile Val	
90 95 100	
ccc atc gag cag gtt ttc cct ccc ggc caa tcc acc gag gaa gtc gaa	451
Pro Ile Glu Gln Val Phe Pro Pro Gly Gln Ser Thr Glu Glu Val Glu	
105 110 115	
gaa tcc aac cgc acc gcg ttc atc tct tcg gag tct tcc gca acg atc	499
Glu Ser Asn Arg Thr Ala Phe Ile Ser Ser Glu Ser Ser Ala Thr Ile	
120 125 130	
gcc gcg atg aat tac ctc aac att ccc gtc gaa gtt gaa gtt gca gaa	547
Ala Ala Met Asn Tyr Leu Asn Ile Pro Val Glu Val Glu Val Ala Glu	
135 140 145	
gtc ctc acc gac agc gcc gca acc gga att ttc gaa ccc ggc gac aaa	595
Val Leu Thr Asp Ser Ala Ala Thr Gly Ile Phe Glu Pro Gly Asp Lys	
150 155 160 165	
ctt ctc agc atc gac ggc acc gca atc tcc act ccc ggc gat gca caa	643
Leu Leu Ser Ile Asp Gly Thr Ala Ile Ser Thr Pro Gly Asp Ala Gln	
170 175 180	
acc atc gtg cga tcg aaa gct ccc ggc gat gag atc acg att tcc tac	691
Thr Ile Val Arg Ser Lys Ala Pro Gly Asp Glu Ile Thr Ile Ser Tyr	
185 190 195	
gag aga aac gat gcg gaa tct caa gca acc atc act ttg agg gaa cac	739
Glu Arg Asn Asp Ala Glu Ser Gln Ala Thr Ile Thr Leu Arg Glu His	
200 205 210	
ccg gat gat tct tcg gtg gcg ttg ttg ggt att tca atg ttg tcg gtg	787
Pro Asp Asp Ser Ser Val Ala Leu Leu Gly Ile Ser Met Leu Ser Val	
215 220 225	
cct tcg agc gcg att gag gtt gat tac aac ttg gaa gat atc ggt ggt	835
Pro Ser Ser Ala Ile Glu Val Asp Tyr Asn Leu Glu Asp Ile Gly Gly	
230 235 240 245	
ccg agc gct ggc atg atg ttt tcg ttg gcg gtc gtc gat aag ctt tcg	883

Pro Ser Ala Gly Met Met Phe Ser Leu Ala Val Val Asp Lys Leu Ser  
 250 255 260

cct ggc gcg ctg aat ggt ggc aag ttt gtc gct ggc act ggc acc atc 931  
 Pro Gly Ala Leu Asn Gly Gly Lys Phe Val Ala Gly Thr Gly Thr Ile  
 265 270 275

gcg gag gac ggg tcg gtg ggc ccg att ggc ggt att gcg cac aag gtg 979  
 Ala Glu Asp Gly Ser Val Gly Pro Ile Gly Gly Ile Ala His Lys Val  
 280 285 290

cgc gct gcg gag gac gcg ggc gcg gaa gtg ttt ttg agc cct gcg gac 1027  
 Arg Ala Ala Glu Asp Ala Gly Ala Glu Val Phe Leu Ser Pro Ala Asp  
 295 300 305

aat tgc gcg gag gcg atg agt gcg aag cct cag gat atg acg atc ttg 1075  
 Asn Cys Ala Glu Ala Met Ser Ala Lys Pro Gln Asp Met Thr Ile Leu  
 310 315 320 325

aag gtg gat tcg ttg tct cag gca atc gat cag atg gct gcc tac aac 1123  
 Lys Val Asp Ser Leu Ser Gln Ala Ile Asp Gln Met Ala Ala Tyr Asn  
 330 335 340

gag ggc tct gat ttc cag acg tgt ggc tagtttttag acgctgaggt 1170  
 Glu Gly Ser Asp Phe Gln Thr Cys Gly  
 345 350

ttc 1173

&lt;210&gt; 40

&lt;211&gt; 350

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 40

Val Asn Arg Arg Ile Lys Thr Leu Thr Trp Gly Ala Ile Pro Leu Val  
 1 5 10 15

Leu Leu Ala Ser Leu Val Ser Ile Asp His Ile Pro Gly Thr Asn Ile  
 20 25 30

Asn Leu Ser Val Pro Tyr Ala Ala Glu Gly Pro Gly Pro Thr Ile Asn  
 35 40 45

Thr Leu Gly Gln Val Asp Gly Glu Asp Val Val Ser Ile Ser Ser Ala  
 50 55 60

Asp Leu Asp Glu Thr Glu Gly Asn Leu Asn Met Thr Thr Val Ser Val  
 65 70 75 80

Arg Ser Gly Met Thr Leu Ser Gln Val Ile Ser Arg Trp Leu Phe Thr  
 85 90 95

Asp Asp Thr Ile Val Pro Ile Glu Gln Val Phe Pro Pro Gly Gln Ser  
 100 105 110

Thr Glu Glu Val Glu Glu Ser Asn Arg Thr Ala Phe Ile Ser Ser Glu  
 115 120 125

Ser Ser Ala Thr Ile Ala Ala Met Asn Tyr Leu Asn Ile Pro Val Glu

130					135					140					
Val	Glu	Val	Ala	Glu	Val	Leu	Thr	Asp	Ser	Ala	Ala	Thr	Gly	Ile	Phe
145					150					155					160
Glu	Pro	Gly	Asp	Lys	Leu	Leu	Ser	Ile	Asp	Gly	Thr	Ala	Ile	Ser	Thr
				165					170					175	
Pro	Gly	Asp	Ala	Gln	Thr	Ile	Val	Arg	Ser	Lys	Ala	Pro	Gly	Asp	Glu
			180					185					190		
Ile	Thr	Ile	Ser	Tyr	Glu	Arg	Asn	Asp	Ala	Glu	Ser	Gln	Ala	Thr	Ile
		195					200					205			
Thr	Leu	Arg	Glu	His	Pro	Asp	Asp	Ser	Ser	Val	Ala	Leu	Leu	Gly	Ile
	210					215					220				
Ser	Met	Leu	Ser	Val	Pro	Ser	Ser	Ala	Ile	Glu	Val	Asp	Tyr	Asn	Leu
225					230					235					240
Glu	Asp	Ile	Gly	Gly	Pro	Ser	Ala	Gly	Met	Met	Phe	Ser	Leu	Ala	Val
				245					250					255	
Val	Asp	Lys	Leu	Ser	Pro	Gly	Ala	Leu	Asn	Gly	Gly	Lys	Phe	Val	Ala
			260					265					270		
Gly	Thr	Gly	Thr	Ile	Ala	Glu	Asp	Gly	Ser	Val	Gly	Pro	Ile	Gly	Gly
		275					280					285			
Ile	Ala	His	Lys	Val	Arg	Ala	Ala	Glu	Asp	Ala	Gly	Ala	Glu	Val	Phe
	290					295					300				
Leu	Ser	Pro	Ala	Asp	Asn	Cys	Ala	Glu	Ala	Met	Ser	Ala	Lys	Pro	Gln
305					310					315					320
Asp	Met	Thr	Ile	Leu	Lys	Val	Asp	Ser	Leu	Ser	Gln	Ala	Ile	Asp	Gln
				325					330					335	
Met	Ala	Ala	Tyr	Asn	Glu	Gly	Ser	Asp	Phe	Gln	Thr	Cys	Gly		
			340					345					350		

<210> 41  
 <211> 1119  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1096)  
 <223> RXA01654

<400> 41  
 cttttgctgt tatgaaagtg aactaaagat atctatgcgc agaccgggtct cacgccgcgc 60  
 catttttgca acatctgttt tggttgcggg ggtgagcatc atg tca cct tcg gcc 115  
 Met Ser Pro Ser Ala  
 1 5  
 aac gca gct gag gct ccg gca tcg gaa tgg gtg aat acg aca gcg atc 163  
 Asn Ala Ala Glu Ala Pro Ala Ser Glu Trp Val Asn Thr Thr Ala Ile

10										15										20									
gta	gat	caa	gcg	aat	gct	cag	ttg	tcg	cag	ttt	ggc	gtg	agt	ctt	gac					211									
Val	Asp	Gln	Ala	Asn	Ala	Gln	Leu	Ser	Gln	Phe	Gly	Val	Ser	Leu	Asp														
			25					30					35																
cga	agt	gca	gca	gaa	ctt	ttt	gat	gat	cag	gca	aac	tcc	caa	att	gat					259									
Arg	Ser	Ala	Ala	Glu	Leu	Phe	Asp	Asp	Gln	Ala	Asn	Ser	Gln	Ile	Asp														
		40					45					50																	
gca	gcg	ctt	tca	ccg	tat	gcc	gat	aag	gtt	cca	acc	tct	ggc	ggc	cag					307									
Ala	Ala	Leu	Ser	Pro	Tyr	Ala	Asp	Lys	Val	Pro	Thr	Ser	Gly	Gly	Gln														
		55				60				65																			
gta	gtc	gag	caa	agt	ctt	cag	gtt	gtg	gag	cag	gaa	gtt	caa	aag	gca					355									
Val	Val	Glu	Gln	Ser	Leu	Gln	Val	Val	Glu	Gln	Glu	Val	Gln	Lys	Ala														
	70				75				80						85														
ctg	ccc	aac	tat	gaa	atc	cgt	acc	gat	ctg	caa	tcc	cag	gtg	atg	ggc					403									
Leu	Pro	Asn	Tyr	Glu	Ile	Arg	Thr	Asp	Leu	Gln	Ser	Gln	Val	Met	Gly														
				90					95					100															
gca	act	cta	gga	gag	gtg	ctg	cac	cga	gtt	cct	gga	tca	tgg	ttt	gat					451									
Ala	Thr	Leu	Gly	Glu	Val	Leu	His	Arg	Val	Pro	Gly	Ser	Trp	Phe	Asp														
			105					110					115																
gcg	cca	gca	gtt	cct	gaa	gaa	tcc	agg	att	gta	gag	gaa	cag	ggc	aaa					499									
Ala	Pro	Ala	Val	Pro	Glu	Glu	Ser	Arg	Ile	Val	Glu	Glu	Gln	Gly	Lys														
		120					125					130																	
tcc	ctg	tat	ggg	ccc	ggc	acc	ccg	atc	tat	ctc	aac	gga	aat	tcc	atg					547									
Ser	Leu	Tyr	Gly	Pro	Gly	Thr	Pro	Ile	Tyr	Leu	Asn	Gly	Asn	Ser	Met														
		135				140					145																		
tgc	acg	ctt	gcg	gtg	act	gga	act	gat	gca	gat	ggg	cgc	aag	atc	ggc					595									
Cys	Thr	Leu	Ala	Val	Thr	Gly	Thr	Asp	Ala	Asp	Gly	Arg	Lys	Ile	Gly														
	150				155					160				165															
atc	act	gca	ggc	cac	tgt	gga	aaa	tcg	ggc	gat	gca	gtc	cgt	tcg	gct					643									
Ile	Thr	Ala	Gly	His	Cys	Gly	Lys	Ser	Gly	Asp	Ala	Val	Arg	Ser	Ala														
				170					175					180															
gac	tcc	ttc	tgg	gtc	ggc	gat	acc	gga	aca	gtg	gtg	tac	aac	gcg	cct					691									
Asp	Ser	Phe	Trp	Val	Gly	Asp	Thr	Gly	Thr	Val	Val	Tyr	Asn	Ala	Pro														
			185					190					195																
aac	gct	gac	tac	tcc	gtg	atc	gag	ttc	ggc	tcc	aat	gca	gag	ttg	agc					739									
Asn	Ala	Asp	Tyr	Ser	Val	Ile	Glu	Phe	Gly	Ser	Asn	Ala	Glu	Leu	Ser														
		200					205					210																	
aat	acc	tac	aac	ggc	gtc	acc	gcg	aat	gct	gtc	ggc	ggc	ggc	gtg	act					787									
Asn	Thr	Tyr	Asn	Gly	Val	Thr	Ala	Asn	Ala	Val	Gly	Gly	Gly	Val	Thr														
		215				220				225																			
aat	ggc	caa	gaa	gta	tgc	aaa	aac	gga	gtt	gct	act	ggc	tac	acc	tgt					835									
Asn	Gly	Gln	Glu	Val	Cys	Lys	Asn	Gly	Val	Ala	Thr	Gly	Tyr	Thr	Cys														
	230				235				240					245															
ggc	ttg	gtg	tgg	act	gct	gat	gag	cgc	atg	acg	atg	tct	cag	gtg	tgt					883									
Gly	Leu	Val	Trp	Thr	Ala	Asp	Glu	Arg	Met	Thr	Met	Ser	Gln	Val	Cys														
			250					255						260															

gcg ggt cgt ggt gat tcg ggt gct ccg ctg att gca gat ggt cgt gtg 931  
 Ala Gly Arg Gly Asp Ser Gly Ala Pro Leu Ile Ala Asp Gly Arg Val  
                   265                                  270                                  275

gtt ggt ctt gta tct ggt ggt gta att cct gat tac aac ctg gca tgc 979  
 Val Gly Leu Val Ser Gly Gly Val Ile Pro Asp Tyr Asn Leu Ala Cys  
                   280                                  285                                  290

gcc act ccg ttg cag gga cct ttc ttc atg cca acg ctg tca gtg aac 1027  
 Ala Thr Pro Leu Gln Gly Pro Phe Phe Met Pro Thr Leu Ser Val Asn  
                   295                                  300                                  305

atg gat act gtc cta act gat ttg gat tcg cag gat ctt ccc ggt cga 1075  
 Met Asp Thr Val Leu Thr Asp Leu Asp Ser Gln Asp Leu Pro Gly Arg  
                   310                                  315                                  320                                  325

ggt ttt cag cca act gct gga tagaatttag aaaatccgcc gtt 1119  
 Gly Phe Gln Pro Thr Ala Gly  
                                   330

<210> 42  
 <211> 332  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 42  
 Met Ser Pro Ser Ala Asn Ala Ala Glu Ala Pro Ala Ser Glu Trp Val  
   1                                  5                                  10                                  15

Asn Thr Thr Ala Ile Val Asp Gln Ala Asn Ala Gln Leu Ser Gln Phe  
                   20                                  25                                  30

Gly Val Ser Leu Asp Arg Ser Ala Ala Glu Leu Phe Asp Asp Gln Ala  
                   35                                  40                                  45

Asn Ser Gln Ile Asp Ala Ala Leu Ser Pro Tyr Ala Asp Lys Val Pro  
                   50                                  55                                  60

Thr Ser Gly Gly Gln Val Val Glu Gln Ser Leu Gln Val Val Glu Gln  
                   65                                  70                                  75                                  80

Glu Val Gln Lys Ala Leu Pro Asn Tyr Glu Ile Arg Thr Asp Leu Gln  
                   85                                  90                                  95

Ser Gln Val Met Gly Ala Thr Leu Gly Glu Val Leu His Arg Val Pro  
                   100                                  105                                  110

Gly Ser Trp Phe Asp Ala Pro Ala Val Pro Glu Glu Ser Arg Ile Val  
                   115                                  120                                  125

Glu Glu Gln Gly Lys Ser Leu Tyr Gly Pro Gly Thr Pro Ile Tyr Leu  
                   130                                  135                                  140

Asn Gly Asn Ser Met Cys Thr Leu Ala Val Thr Gly Thr Asp Ala Asp  
                   145                                  150                                  155                                  160

Gly Arg Lys Ile Gly Ile Thr Ala Gly His Cys Gly Lys Ser Gly Asp  
                   165                                  170                                  175

Ala Val Arg Ser Ala Asp Ser Phe Trp Val Gly Asp Thr Gly Thr Val  
 180 185 190

Val Tyr Asn Ala Pro Asn Ala Asp Tyr Ser Val Ile Glu Phe Gly Ser  
 195 200 205

Asn Ala Glu Leu Ser Asn Thr Tyr Asn Gly Val Thr Ala Asn Ala Val  
 210 215 220

Gly Gly Gly Val Thr Asn Gly Gln Glu Val Cys Lys Asn Gly Val Ala  
 225 230 235 240

Thr Gly Tyr Thr Cys Gly Leu Val Trp Thr Ala Asp Glu Arg Met Thr  
 245 250 255

Met Ser Gln Val Cys Ala Gly Arg Gly Asp Ser Gly Ala Pro Leu Ile  
 260 265 270

Ala Asp Gly Arg Val Val Gly Leu Val Ser Gly Gly Val Ile Pro Asp  
 275 280 285

Tyr Asn Leu Ala Cys Ala Thr Pro Leu Gln Gly Pro Phe Phe Met Pro  
 290 295 300

Thr Leu Ser Val Asn Met Asp Thr Val Leu Thr Asp Leu Asp Ser Gln  
 305 310 315 320

Asp Leu Pro Gly Arg Gly Phe Gln Pro Thr Ala Gly  
 325 330

<210> 43  
 <211> 2049  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(2026)  
 <223> RXN01868

<400> 43  
 tgacaggcta ccttctgggg tggacatgat ccccaacgct caaccactt gtggcaccaa 60  
 ccacaaaccc tgtggcggtta aatccctag agtaggccac atg aag gat ctt tat 115  
 Met Lys Asp Leu Tyr  
 1 5

cgc ttt gtc aat ggc ctg tgg ctt gac acc cac atc att ccc gac gat 163  
 Arg Phe Val Asn Gly Leu Trp Leu Asp Thr His Ile Ile Pro Asp Asp  
 10 15 20

cgc gcg gtg gac ggc acg ttc cac aag ctg cgc gat gat gct gaa gaa 211  
 Arg Ala Val Asp Gly Thr Phe His Lys Leu Arg Asp Asp Ala Glu Glu  
 25 30 35

gac gtc cat gag atc gtc aag gaa gac act gga cgc gca ggc aca ctt 259  
 Asp Val His Glu Ile Val Lys Glu Asp Thr Gly Arg Ala Gly Thr Leu  
 40 45 50

tat gcc tca ttt atg gat act gac gcc atc aac gct gct ggt gtt gca 307

Tyr	Ala	Ser	Phe	Met	Asp	Thr	Asp	Ala	Ile	Asn	Ala	Ala	Gly	Val	Ala		
55						60					65						
cgc	ctc	gat	gcg	gat	ctg	aac	agg	ctg	tct	gtt	gct	aac	tca	tcg	ttt	355	
Pro	Leu	Asp	Ala	Asp	Leu	Asn	Arg	Leu	Ser	Val	Ala	Asn	Ser	Ser	Phe		
70					75					80					85		
ttc	gca	gct	gct	ctc	ggc	gaa	ctg	gac	cgt	gaa	ggc	gtt	ggc	gcg	cca	403	
Phe	Ala	Ala	Ala	Leu	Gly	Glu	Leu	Asp	Arg	Glu	Gly	Val	Gly	Ala	Pro		
				90					95						100		
gta	ggt	ttc	tgg	gtg	gag	aag	gat	tct	tcc	tcc	aac	gaa	tcc	gtc	gcc	451	
Val	Gly	Phe	Trp	Val	Glu	Lys	Asp	Ser	Ser	Ser	Asn	Glu	Ser	Val	Ala		
			105					110						115			
tat	gtc	atc	cag	tcc	ggc	ctc	ggc	ctg	ccc	gat	gag	gct	tat	tac	cgc	499	
Tyr	Val	Ile	Gln	Ser	Gly	Leu	Gly	Leu	Pro	Asp	Glu	Ala	Tyr	Tyr	Arg		
			120					125									
gag	gag	gca	cac	gcc	gaa	act	ctc	gcg	gcc	tac	aaa	gag	cac	gtt	gag	547	
Glu	Glu	Ala	His	Ala	Glu	Thr	Leu	Ala	Ala	Tyr	Lys	Glu	His	Val	Glu		
			135				140				145						
cgc	atg	ctc	ggc	tac	ttg	gat	aac	agc	cgc	ctc	ttc	ggt	ctg	tcg	gct	595	
Arg	Met	Leu	Gly	Tyr	Leu	Asp	Asn	Ser	Arg	Leu	Phe	Gly	Leu	Ser	Ala		
150					155					160					165		
gct	tcc	gct	gcc	gca	cga	att	gtc	gcc	ctg	gaa	acg	gaa	atc	gct	gct	643	
Ala	Ser	Ala	Ala	Ala	Arg	Ile	Val	Ala	Leu	Glu	Thr	Glu	Ile	Ala	Ala		
				170					175						180		
ggc	cac	tgg	gat	gtc	gtg	aag	acc	cgc	gac	gcc	gta	gcc	acc	tac	aac	691	
Gly	His	Trp	Asp	Val	Val	Lys	Thr	Arg	Asp	Ala	Val	Ala	Thr	Tyr	Asn		
			185					190						195			
ccc	acc	gaa	ctc	ggc	gcg	ctg	cca	cca	aag	gtc	cgc	acg	ctg	ctc	agt	739	
Pro	Thr		Leu	Gly	Ala	Leu	Pro	Pro	Lys	Val	Arg	Thr	Leu	Leu	Ser		
			200				205										
tcc	gca	ggc	ctc	ccg	gac	cag	cgc	ctg	gta	tcg	atg	atg	ccg	tca	tac	787	
Ser	Ala	Gly	Leu	Pro	Asp	Gln	Arg	Leu	Val	Ser	Met	Met	Pro	Ser	Tyr		
			215				220					225					
ctc	gac	cac	ctc	aac	ggc	ttg	ctt	gtc	gac	gac	cgc	ctc	ccc	gat	tgg	835	
Leu	Asp	His	Leu	Asn	Gly	Leu	Leu	Val	Asp	Asp	Arg	Leu	Pro	Asp	Trp		
230					235					240					245		
cag	cta	tgg	gca	acc	tgg	cac	atc	ttg	agg	tct	cga	gca	gga	ctg	ttg	883	
Gln	Leu	Trp	Ala	Thr	Trp	His	Ile	Leu	Arg	Ser	Arg	Ala	Gly	Leu	Leu		
				250					255						260		
acc	gag	gaa	att	agc	caa	gca	aac	ttc	gac	ttc	tat	ggc	acc	aaa	ctg	931	
Thr	Glu	Glu	Ile	Ser	Gln	Ala	Asn	Phe	Asp	Phe	Tyr	Gly	Thr	Lys	Leu		
			265					270						275			
tcc	ggc	gcc	acc	gag	caa	aaa	gat	cga	tgg	aag	cgt	gct	gtc	ggc	ctg	979	
Ser	Gly	Ala	Thr	Glu	Gln	Lys	Asp	Arg	Trp	Lys	Arg	Ala	Val	Gly	Leu		
			280				285							290			
gca	gag	cgc	atg	gtg	ggc	gag	gaa	atc	ggg	caa	cga	ttc	gtc	gaa	agg	1027	
Ala	Glu	Arg	Met	Val	Gly	Glu	Glu	Ile	Gly	Gln	Arg	Phe	Val	Glu	Arg		

295	300	305	
cat ttt cct gca agc tcc aag gag cac atg ctt gag ctc gtc gac tac His Phe Pro Ala Ser Ser Lys Glu His Met Leu Glu Leu Val Asp Tyr 310 315 320 325			1075
ctg gtt gcc gcc tac cgt gat cgc att tcc aac ctc gaa tgg atg acg Leu Val Ala Ala Tyr Arg Asp Arg Ile Ser Asn Leu Glu Trp Met Thr 330 335 340			1123
ccc gcc acc cgc gag cgt gcc ctg gaa aag ttg ggc aaa ttc aac gcg Pro Ala Thr Arg Glu Arg Ala Leu Glu Lys Leu Gly Lys Phe Asn Ala 345 350 355			1171
aaa atc ggc tac ccc gac aag tgg cgc tcc tac gaa ggc ctc gaa ttc Lys Ile Gly Tyr Pro Asp Lys Trp Arg Ser Tyr Glu Gly Leu Glu Phe 360 365 370			1219
ggc tcc gac ctg gtg gac aac tcc cgc aag ggc tcc gcg ttc ctc cat Gly Ser Asp Leu Val Asp Asn Ser Arg Lys Gly Ser Ala Phe Leu His 375 380 385			1267
gac tat gag ctg ggc aag atc ggc aaa cca gcc gac cgc gac gaa tgg Asp Tyr Glu Leu Gly Lys Ile Gly Lys Pro Ala Asp Arg Asp Glu Trp 390 395 400 405			1315
gtc acc acc cca caa acc gtc aac gcc ttc tac aac ccc gtg gtc aac Val Thr Thr Pro Gln Thr Val Asn Ala Phe Tyr Asn Pro Val Val Asn 410 415 420			1363
gac atc acc ttc ccc gca gcc atc ctg cgc gca cca ttc ttc gac ccc Asp Ile Thr Phe Pro Ala Ala Ile Leu Arg Ala Pro Phe Phe Asp Pro 425 430 435			1411
gaa gca gaa gcc gca gaa aac ttc ggt gca atc ggt gct gtg atc gga Glu Ala Glu Ala Ala Glu Asn Phe Gly Ala Ile Gly Ala Val Ile Gly 440 445 450			1459
cac gaa atc ggc cac ggc ttt gac gat caa ggc agc caa tac gac ggc His Glu Ile Gly His Gly Phe Asp Asp Gln Gly Ser Gln Tyr Asp Gly 455 460 465			1507
gac ggc aac ctc aac tcc tgg tgg acc gac gaa gac cgc tcc gca ttc Asp Gly Asn Leu Asn Ser Trp Trp Thr Asp Glu Asp Arg Ser Ala Phe 470 475 480 485			1555
gag cag ctc acc tca cgt ctg gtc acc caa ttc agc gga ctc gtc cct Glu Gln Leu Thr Ser Arg Leu Val Thr Gln Phe Ser Gly Leu Val Pro 490 495 500			1603
gcc gtc ctg acc tct gaa gga atc gac acc gac ggc gtc aac ggt gaa Ala Val Leu Thr Ser Glu Gly Ile Asp Thr Asp Gly Val Asn Gly Glu 505 510 515			1651
ttc act ctc ggc gaa aac atc ggt gac ctc ggc gga ttg ggc atc gct Phe Thr Leu Gly Glu Asn Ile Gly Asp Leu Gly Gly Leu Gly Ile Ala 520 525 530			1699
gtc gtt gcc tac gaa aag tac ctc gca gac cgt ggc caa acc ttt gaa Val Val Ala Tyr Glu Lys Tyr Leu Ala Asp Arg Gly Gln Thr Phe Glu 535 540 545			1747



acc tca cca gtc caa aaa ttc gaa gca gaa ggc gcc gag gaa ggc ctg 1795  
 Thr Ser Pro Val Gln Lys Phe Glu Ala Glu Gly Ala Glu Glu Gly Leu  
 550 555 560 565

gcc gag caa gaa ttc aac ggt ctc caa cgc ctc ttc ctg tcc tgg gct 1843  
 Ala Glu Gln Glu Phe Asn Gly Leu Gln Arg Leu Phe Leu Ser Trp Ala  
 570 575 580

cgc gtg tgg cgc acc aaa atc cgc cca cag atg gcc gtc caa tac ctg 1891  
 Arg Val Trp Arg Thr Lys Ile Arg Pro Gln Met Ala Val Gln Tyr Leu  
 585 590 595

gcc atc gac cca cac tcc cct gca gaa ttc cgc tgc aat gtc atc gcc 1939  
 Ala Ile Asp Pro His Ser Pro Ala Glu Phe Arg Cys Asn Val Ile Ala  
 600 605 610

gga aac gtc gct gaa ttc tac gaa gca ttc gac gtc ccc gaa gat gca 1987  
 Gly Asn Val Ala Glu Phe Tyr Glu Ala Phe Asp Val Pro Glu Asp Ala  
 615 620 625

cct gtg tac atc aag cca gaa gag cgc cta gct atc tgg tagttgtag 2036  
 Pro Val Tyr Ile Lys Pro Glu Glu Arg Leu Ala Ile Trp  
 630 635 640

ttggtattga aaa 2049

<210> 44  
 <211> 642  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 44  
 Met Lys Asp Leu Tyr Arg Phe Val Asn Gly Leu Trp Leu Asp Thr His  
 1 5 10 15

Ile Ile Pro Asp Asp Arg Ala Val Asp Gly Thr Phe His Lys Leu Arg  
 20 25 30

Asp Asp Ala Glu Glu Asp Val His Glu Ile Val Lys Glu Asp Thr Gly  
 35 40 45

Arg Ala Gly Thr Leu Tyr Ala Ser Phe Met Asp Thr Asp Ala Ile Asn  
 50 55 60

Ala Ala Gly Val Ala Pro Leu Asp Ala Asp Leu Asn Arg Leu Ser Val  
 65 70 75 80

Ala Asn Ser Ser Phe Phe Ala Ala Ala Leu Gly Glu Leu Asp Arg Glu  
 85 90 95

Gly Val Gly Ala Pro Val Gly Phe Trp Val Glu Lys Asp Ser Ser Ser  
 100 105 110

Asn Glu Ser Val Ala Tyr Val Ile Gln Ser Gly Leu Gly Leu Pro Asp  
 115 120 125

Glu Ala Tyr Tyr Arg Glu Glu Ala His Ala Glu Thr Leu Ala Ala Tyr  
 130 135 140

Lys Glu His Val Glu Arg Met Leu Gly Tyr Leu Asp Asn Ser Arg Leu  
 145 150 155 160  
 Phe Gly Leu Ser Ala Ala Ser Ala Ala Ala Arg Ile Val Ala Leu Glu  
 165 170 175  
 Thr Glu Ile Ala Ala Gly His Trp Asp Val Val Lys Thr Arg Asp Ala  
 180 185 190  
 Val Ala Thr Tyr Asn Pro Thr Glu Leu Gly Ala Leu Pro Pro Lys Val  
 195 200 205  
 Arg Thr Leu Leu Ser Ser Ala Gly Leu Pro Asp Gln Arg Leu Val Ser  
 210 215 220  
 Met Met Pro Ser Tyr Leu Asp His Leu Asn Gly Leu Leu Val Asp Asp  
 225 230 235 240  
 Arg Leu Pro Asp Trp Gln Leu Trp Ala Thr Trp His Ile Leu Arg Ser  
 245 250 255  
 Arg Ala Gly Leu Leu Thr Glu Glu Ile Ser Gln Ala Asn Phe Asp Phe  
 260 265 270  
 Tyr Gly Thr Lys Leu Ser Gly Ala Thr Glu Gln Lys Asp Arg Trp Lys  
 275 280 285  
 Arg Ala Val Gly Leu Ala Glu Arg Met Val Gly Glu Glu Ile Gly Gln  
 290 295 300  
 Arg Phe Val Glu Arg His Phe Pro Ala Ser Ser Lys Glu His Met Leu  
 305 310 315 320  
 Glu Leu Val Asp Tyr Leu Val Ala Ala Tyr Arg Asp Arg Ile Ser Asn  
 325 330 335  
 Leu Glu Trp Met Thr Pro Ala Thr Arg Glu Arg Ala Leu Glu Lys Leu  
 340 345 350  
 Gly Lys Phe Asn Ala Lys Ile Gly Tyr Pro Asp Lys Trp Arg Ser Tyr  
 355 360 365  
 Glu Gly Leu Glu Phe Gly Ser Asp Leu Val Asp Asn Ser Arg Lys Gly  
 370 375 380  
 Ser Ala Phe Leu His Asp Tyr Glu Leu Gly Lys Ile Gly Lys Pro Ala  
 385 390 395 400  
 Asp Arg Asp Glu Trp Val Thr Thr Pro Gln Thr Val Asn Ala Phe Tyr  
 405 410 415  
 Asn Pro Val Val Asn Asp Ile Thr Phe Pro Ala Ala Ile Leu Arg Ala  
 420 425 430  
 Pro Phe Phe Asp Pro Glu Ala Glu Ala Ala Glu Asn Phe Gly Ala Ile  
 435 440 445  
 Gly Ala Val Ile Gly His Glu Ile Gly His Gly Phe Asp Asp Gln Gly  
 450 455 460  
 Ser Gln Tyr Asp Gly Asp Gly Asn Leu Asn Ser Trp Trp Thr Asp Glu

```
<210> 45
<211> 1734
<212> DNA
<213> Corynebacterium glutamicum
```

```
<220>
<221> CDS
<222> (101)..(1711)
<223> FRXA01868
```

```

<400> 45
atgcggatct gaacaggctg tctgttgcta actcatcggt ttcgcagctg ctctcggcga 60

actggaccgt gaaggcggtt gcgcgccagt aggtttctgg gtg gag aag gat tct 115
Val Glu Lys Asp Ser
1 5

tcc tcc aac gaa tcc gtc gcc tat gtc atc cag tcc ggc ctc ggc ctg 163
Ser Ser Asn Glu Ser Val Ala Tyr Val Ile Gln Ser Gly Leu Gly Leu
10 15 20

ccc gat gag gct tat tac cgc gag gag gca cac gcc gaa act ctc gcg 211
Pro Asp Glu Ala Tyr Tyr Arg Glu Glu Ala His Ala Glu Thr Leu Ala
25 30 35

```

gcc tac aaa gag cac gtt gag cgc atg ctc ggc tac ttg gat aac agc	259
Ala Tyr Lys Glu His Val Glu Arg Met Leu Gly Tyr Leu Asp Asn Ser	
40 45 50	
cgc ctc ttc ggt ctg tgc gct gct tcc gct gcc gca cga att gtc gcc	307
Arg Leu Phe Gly Leu Ser Ala Ala Ser Ala Ala Ala Arg Ile Val Ala	
55 60 65	
ctg gaa acg gaa atc gct gct ggc cac tgg gat gtc gtg aag acc cgc	355
Leu Glu Thr Glu Ile Ala Ala Gly His Trp Asp Val Val Lys Thr Arg	
70 75 80 85	
gac gcc gta gcc acc tac aac ccc acc gaa ctc ggc gcg ctg cca cca	403
Asp Ala Val Ala Thr Tyr Asn Pro Thr Glu Leu Gly Ala Leu Pro Pro	
90 95 100	
aag gtc cgc acg ctg ctc agt tcc gca ggc ctc ccg gac cag cgc ctg	451
Lys Val Arg Thr Leu Leu Ser Ser Ala Gly Leu Pro Asp Gln Arg Leu	
105 110 115	
gta tgc atg atg ccg tca tac ctc gac cac ctc aac ggc ttg ctt gtc	499
Val Ser Met Met Pro Ser Tyr Leu Asp His Leu Asn Gly Leu Leu Val	
120 125 130	
gac gac cgc ctc ccc gat tgg cag cta tgg gca acc tgg cac atc ttg	547
Asp Asp Arg Leu Pro Asp Trp Gln Leu Trp Ala Thr Trp His Ile Leu	
135 140 145	
agg tct cga gca gga ctg ttg acc gag gaa att agc caa gca aac ttc	595
Arg Ser Arg Ala Gly Leu Leu Thr Glu Glu Ile Ser Gln Ala Asn Phe	
150 155 160 165	
gac ttc tat ggc acc aaa ctg tcc ggc gcc acc gag caa aaa gat cga	643
Asp Phe Tyr Gly Thr Lys Leu Ser Gly Ala Thr Glu Gln Lys Asp Arg	
170 175 180	
tgg aag cgt gct gtc ggc ctg gca gag cgc atg gtg ggc gag gaa atc	691
Trp Lys Arg Ala Val Gly Leu Ala Glu Arg Met Val Gly Glu Glu Ile	
185 190 195	
ggg caa cga ttc gtc gaa agg cat ttt cct gca agc tcc aag gag cac	739
Gly Gln Arg Phe Val Glu Arg His Phe Pro Ala Ser Ser Lys Glu His	
200 205 210	
atg ctt gag ctc gtc gac tac ctg gtt gcc gcc tac cgt gat cgc att	787
Met Leu Glu Leu Val Asp Tyr Leu Val Ala Ala Tyr Arg Asp Arg Ile	
215 220 225	
tcc aac ctc gaa tgg atg acg ccc gcc acc cgc gag cgt gcc ctg gaa	835
Ser Asn Leu Glu Trp Met Thr Pro Ala Thr Arg Glu Arg Ala Leu Glu	
230 235 240 245	
aag ttg ggc aaa ttc aac gcg aaa atc ggc tac ccc gac aag tgg cgc	883
Lys Leu Gly Lys Phe Asn Ala Lys Ile Gly Tyr Pro Asp Lys Trp Arg	
250 255 260	
tcc tac gaa ggc ctc gaa ttc ggc tcc gac ctg gtg gac aac tcc cgc	931
Ser Tyr Glu Gly Leu Glu Phe Gly Ser Asp Leu Val Asp Asn Ser Arg	
265 270 275	
aag ggc tcc gcg ttc ctc cat gac tat gag ctg ggc aag atc ggc aaa	979

Lys Gly Ser Ala Phe Leu His Asp Tyr Glu Leu Gly Lys Ile Gly Lys	
280	285
cca gcc gac cgc gac gaa tgg gtc acc acc cca caa acc gtc aac gcc	1027
Pro Ala Asp Arg Asp Glu Trp Val Thr Thr Pro Gln Thr Val Asn Ala	
295	300
305	
ttc tac aac ccc gtg gtc aac gac atc acc ttc ccc gca gcc atc ctg	1075
Phe Tyr Asn Pro Val Val Asn Asp Ile Thr Phe Pro Ala Ala Ile Leu	
310	315
320	325
cgc gca cca ttc ttc gac ccc gaa gca gaa gcc gca gaa aac ttc ggt	1123
Arg Ala Pro Phe Phe Asp Pro Glu Ala Glu Ala Ala Glu Asn Phe Gly	
330	335
340	
gca atc ggt gct gtg atc gga cac gaa atc ggc cac ggc ttt gac gat	1171
Ala Ile Gly Ala Val Ile Gly His Glu Ile Gly His Gly Phe Asp Asp	
345	350
355	
caa ggc agc caa tac gac ggc gac ggc aac ctc aac tcc tgg tgg acc	1219
Gln Gly Ser Gln Tyr Asp Gly Asp Gly Asn Leu Asn Ser Trp Trp Thr	
360	365
370	
gac gaa gac cgc tcc gca ttc gag cag ctc acc tca cgt ctg gtc acc	1267
Asp Glu Asp Arg Ser Ala Phe Glu Gln Leu Thr Ser Arg Leu Val Thr	
375	380
385	
caa ttc agc gga ctc gtc cct gcc gtc ctg acc tct gaa gga atc gac	1315
Gln Phe Ser Gly Leu Val Pro Ala Val Leu Thr Ser Glu Gly Ile Asp	
390	395
400	405
acc gac ggc gtc aac ggt gaa ttc act ctc ggc gaa aac atc ggt gac	1363
Thr Asp Gly Val Asn Gly Glu Phe Thr Leu Gly Glu Asn Ile Gly Asp	
410	415
420	
ctc ggc gga ttg ggc atc gct gtc gtt gcc tac gaa aag tac ctc gca	1411
Leu Gly Gly Leu Gly Ile Ala Val Val Ala Tyr Glu Lys Tyr Leu Ala	
425	430
435	
gac cgt ggc caa acc ttt gaa acc tca cca gtc caa aaa ttc gaa gca	1459
Asp Arg Gly Gln Thr Phe Glu Thr Ser Pro Val Gln Lys Phe Glu Ala	
440	445
450	
gaa ggc gcc gag gaa ggc ctg gcc gag caa gaa ttc aac ggt ctc caa	1507
Glu Gly Ala Glu Glu Gly Leu Ala Glu Gln Glu Phe Asn Gly Leu Gln	
455	460
465	
cgc ctc ttc ctg tcc tgg gct cgc gtg tgg cgc acc aaa atc cgc cca	1555
Arg Leu Phe Leu Ser Trp Ala Arg Val Trp Arg Thr Lys Ile Arg Pro	
470	475
480	485
cag atg gcc gtc caa tac ctg gcc atc gac cca cac tcc cct gca gaa	1603
Gln Met Ala Val Gln Tyr Leu Ala Ile Asp Pro His Ser Pro Ala Glu	
490	495
500	
ttc cgc tgc aat gtc atc gcc gga aac gtc gct gaa ttc tac gaa gca	1651
Phe Arg Cys Asn Val Ile Ala Gly Asn Val Ala Glu Phe Tyr Glu Ala	
505	510
515	
ttc gac gtc ccc gaa gat gca cct gtg tac atc aag cca gaa gag cgc	1699
Phe Asp Val Pro Glu Asp Ala Pro Val Tyr Ile Lys Pro Glu Glu Arg	

520                                      525                                      530  
 cta gct atc tgg tagttgtag ttggtattga aaa                                      1734  
 Leu Ala Ile Trp  
 535  
  
 <210> 46  
 <211> 537  
 <212> PRT  
 <213> *Corynebacterium glutamicum*  
  
 <400> 46  
 Val Glu Lys Asp Ser Ser Ser Asn Glu Ser Val Ala Tyr Val Ile Gln  
 1                                      5                                      10                                      15  
 Ser Gly Leu Gly Leu Pro Asp Glu Ala Tyr Tyr Arg Glu Glu Ala His  
                                     20                                      25                                      30  
 Ala Glu Thr Leu Ala Ala Tyr Lys Glu His Val Glu Arg Met Leu Gly  
                                     35                                      40                                      45  
 Tyr Leu Asp Asn Ser Arg Leu Phe Gly Leu Ser Ala Ala Ser Ala Ala  
                                     50                                      55                                      60  
 Ala Arg Ile Val Ala Leu Glu Thr Glu Ile Ala Ala Gly His Trp Asp  
                                     65                                      70                                      75                                      80  
 Val Val Lys Thr Arg Asp Ala Val Ala Thr Tyr Asn Pro Thr Glu Leu  
                                     85                                      90                                      95  
 Gly Ala Leu Pro Pro Lys Val Arg Thr Leu Leu Ser Ser Ala Gly Leu  
                                     100                                      105                                      110  
 Pro Asp Gln Arg Leu Val Ser Met Met Pro Ser Tyr Leu Asp His Leu  
                                     115                                      120                                      125  
 Asn Gly Leu Leu Val Asp Asp Arg Leu Pro Asp Trp Gln Leu Trp Ala  
                                     130                                      135                                      140  
 Thr Trp His Ile Leu Arg Ser Arg Ala Gly Leu Leu Thr Glu Glu Ile  
                                     145                                      150                                      155                                      160  
 Ser Gln Ala Asn Phe Asp Phe Tyr Gly Thr Lys Leu Ser Gly Ala Thr  
                                     165                                      170                                      175  
 Glu Gln Lys Asp Arg Trp Lys Arg Ala Val Gly Leu Ala Glu Arg Met  
                                     180                                      185                                      190  
 Val Gly Glu Glu Ile Gly Gln Arg Phe Val Glu Arg His Phe Pro Ala  
                                     195                                      200                                      205  
 Ser Ser Lys Glu His Met Leu Glu Leu Val Asp Tyr Leu Val Ala Ala  
                                     210                                      215                                      220  
 Tyr Arg Asp Arg Ile Ser Asn Leu Glu Trp Met Thr Pro Ala Thr Arg  
                                     225                                      230                                      235                                      240  
 Glu Arg Ala Leu Glu Lys Leu Gly Lys Phe Asn Ala Lys Ile Gly Tyr  
                                     245                                      250                                      255

Pro Asp Lys Trp Arg Ser Tyr Glu Gly Leu Glu Phe Gly Ser Asp Leu  
 260 265 270  
 Val Asp Asn Ser Arg Lys Gly Ser Ala Phe Leu His Asp Tyr Glu Leu  
 275 280 285  
 Gly Lys Ile Gly Lys Pro Ala Asp Arg Asp Glu Trp Val Thr Thr Pro  
 290 295 300  
 Gln Thr Val Asn Ala Phe Tyr Asn Pro Val Val Asn Asp Ile Thr Phe  
 305 310 315 320  
 Pro Ala Ala Ile Leu Arg Ala Pro Phe Phe Asp Pro Glu Ala Glu Ala  
 325 330 335  
 Ala Glu Asn Phe Gly Ala Ile Gly Ala Val Ile Gly His Glu Ile Gly  
 340 345 350  
 His Gly Phe Asp Asp Gln Gly Ser Gln Tyr Asp Gly Asp Gly Asn Leu  
 355 360 365  
 Asn Ser Trp Trp Thr Asp Glu Asp Arg Ser Ala Phe Glu Gln Leu Thr  
 370 375 380  
 Ser Arg Leu Val Thr Gln Phe Ser Gly Leu Val Pro Ala Val Leu Thr  
 385 390 395 400  
 Ser Glu Gly Ile Asp Thr Asp Gly Val Asn Gly Glu Phe Thr Leu Gly  
 405 410 415  
 Glu Asn Ile Gly Asp Leu Gly Gly Leu Gly Ile Ala Val Val Ala Tyr  
 420 425 430  
 Glu Lys Tyr Leu Ala Asp Arg Gly Gln Thr Phe Glu Thr Ser Pro Val  
 435 440 445  
 Gln Lys Phe Glu Ala Glu Gly Ala Glu Glu Gly Leu Ala Glu Gln Glu  
 450 455 460  
 Phe Asn Gly Leu Gln Arg Leu Phe Leu Ser Trp Ala Arg Val Trp Arg  
 465 470 475 480  
 Thr Lys Ile Arg Pro Gln Met Ala Val Gln Tyr Leu Ala Ile Asp Pro  
 485 490 495  
 His Ser Pro Ala Glu Phe Arg Cys Asn Val Ile Ala Gly Asn Val Ala  
 500 505 510  
 Glu Phe Tyr Glu Ala Phe Asp Val Pro Glu Asp Ala Pro Val Tyr Ile  
 515 520 525  
 Lys Pro Glu Glu Arg Leu Ala Ile Trp  
 530 535

<210> 47  
 <211> 426  
 <212> DNA  
 <213> Corynebacterium glutamicum  
 <220>

&lt;221&gt; CDS

&lt;222&gt; (101)..(403)

&lt;223&gt; FRXA01869

&lt;400&gt; 47

tgacaggcta ccttctgggg tggacatgat ccccaacgct caaccactt gtggcaccaa 60

ccacaaaccc tgtggcggtta aatcccctag agtaggccac atg aag gat ctt tat 115  
Met Lys Asp Leu Tyr  
1 5

cgc ttt gtc aat ggc ctg tgg ctt gac acc cac atc att ccc gac gat 163  
Arg Phe Val Asn Gly Leu Trp Leu Asp Thr His Ile Ile Pro Asp Asp  
10 15 20

cgc gcg gtg gac ggc acg ttc cac aag ctg cgc gat gat gct gaa gaa 211  
Arg Ala Val Asp Gly Thr Phe His Lys Leu Arg Asp Asp Ala Glu Glu  
25 30 35

gac gtc cat gag atc gtc aag gaa gac act gga cgc gca ggc aca ctt 259  
Asp Val His Glu Ile Val Lys Glu Asp Thr Gly Arg Ala Gly Thr Leu  
40 45 50

tat gcc tca ttt atg gat act gac gcc atc aac gct gct ggt gtt gca 307  
Tyr Ala Ser Phe Met Asp Thr Asp Ala Ile Asn Ala Ala Gly Val Ala  
55 60 65

cgc ctc gat gcg gat ctg aac agg ctg tct gtt gct aac tca tcg ttt 355  
Pro Leu Asp Ala Asp Leu Asn Arg Leu Ser Val Ala Asn Ser Ser Phe  
70 75 80 85

tcg cag ctg ctc tcg gcg aac tgg acc gtg aag gcg ttg gcg cgc cag 403  
Ser Gln Leu Leu Ser Ala Asn Trp Thr Val Lys Ala Leu Ala Arg Gln  
90 95 100

taggtttctg ggtggagaag gat 426

&lt;210&gt; 48

&lt;211&gt; 101

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 48

Met Lys Asp Leu Tyr Arg Phe Val Asn Gly Leu Trp Leu Asp Thr His  
1 5 10 15

Ile Ile Pro Asp Asp Arg Ala Val Asp Gly Thr Phe His Lys Leu Arg  
20 25 30

Asp Asp Ala Glu Glu Asp Val His Glu Ile Val Lys Glu Asp Thr Gly  
35 40 45

Arg Ala Gly Thr Leu Tyr Ala Ser Phe Met Asp Thr Asp Ala Ile Asn  
50 55 60

Ala Ala Gly Val Ala Pro Leu Asp Ala Asp Leu Asn Arg Leu Ser Val  
65 70 75 80

Ala Asn Ser Ser Phe Ser Gln Leu Leu Ser Ala Asn Trp Thr Val Lys  
85 90 95



Ala Leu Ala Arg Gln  
100

<210> 49  
<211> 2898  
<212> DNA  
<213> *Corynebacterium glutamicum*

<220>  
<221> CDS  
<222> (101)..(2875)  
<223> RXN03028

<400> 49  
ctaattccaa gccgagctga aaaagtctgg aagttttgcc caataagggc gttaaagtgg 60  
gtgaaagcga atttagaaat aaagaattaa ggggagagac atg ttc gag agg ttt 115  
Met Phe Glu Arg Phe  
1 5  
acc gat cgt gca cgc cgc gtg att gtg ctc gcg cag gaa gag gcg cgc 163  
Thr Asp Arg Ala Arg Arg Val Ile Val Leu Ala Gln Glu Glu Ala Arg  
10 15 20  
atg ctc aac cac aat tac atc ggc acg gag cac att ctc ctc ggc ctc 211  
Met Leu Asn His Asn Tyr Ile Gly Thr Glu His Ile Leu Leu Gly Leu  
25 30 35  
att cac gag ggc gag ggc gtt gca gcc aag gct ttg gaa tcc atg gga 259  
Ile His Glu Gly Glu Gly Val Ala Ala Lys Ala Leu Glu Ser Met Gly  
40 45 50  
att tcc ctg gac gcc gtc cgc cag gaa gtc gaa gag att atc ggc cag 307  
Ile Ser Leu Asp Ala Val Arg Gln Glu Val Glu Glu Ile Ile Gly Gln  
55 60 65  
ggc tcc cag ccc acc acc ggc cac att cct ttt act cca cgt gcc aag 355  
Gly Ser Gln Pro Thr Thr Gly His Ile Pro Phe Thr Pro Arg Ala Lys  
70 75 80 85  
aag gtc ctg gag ctc agc ctc cgc gaa ggc cta caa atg gga cac aag 403  
Lys Val Leu Glu Leu Ser Leu Arg Glu Gly Leu Gln Met Gly His Lys  
90 95 100  
tac atc ggt act gag ttc ctg ctt ctc ggt ttg atc cgt gag ggc gag 451  
Tyr Ile Gly Thr Glu Phe Leu Leu Leu Gly Leu Ile Arg Glu Gly Glu  
105 110 115  
ggc gtt gct gcc cag gtc ctg gtc aag ctt ggt gct gat ctg cca cgc 499  
Gly Val Ala Ala Gln Val Leu Val Lys Leu Gly Ala Asp Leu Pro Arg  
120 125 130  
gtg cgt cag caa gtt att cag ctt ctc tcc ggc tac gaa ggt ggc cag 547  
Val Arg Gln Gln Val Ile Gln Leu Leu Ser Gly Tyr Glu Gly Gly Gln  
135 140 145  
ggc gga tcc cca gag ggc ggc cag ggc gcc cct act ggc ggt gac gct 595  
Gly Gly Ser Pro Glu Gly Gly Gln Gly Ala Pro Thr Gly Gly Asp Ala  
150 155 160 165

gtt ggt gca gga gct gct cct ggc ggt cgt cca tct tcg ggc agc cca	643
Val Gly Ala Gly Ala Ala Pro Gly Gly Arg Pro Ser Ser Gly Ser Pro	
170 175 180	
ggc gag cgt tct acc tct ttg gtc ctt gac cag ttc ggc cgc aac ctc	691
Gly Glu Arg Ser Thr Ser Leu Val Leu Asp Gln Phe Gly Arg Asn Leu	
185 190 195	
acc cag gct gca aag gac ggc aag ctg gat cca gtt gtt ggt cgc gat	739
Thr Gln Ala Ala Lys Asp Gly Lys Leu Asp Pro Val Val Gly Arg Asp	
200 205 210	
aag gaa atc gag cgc atc atg cag gtg ctc tcc cgt cgt acc aag aac	787
Lys Glu Ile Glu Arg Ile Met Gln Val Leu Ser Arg Arg Thr Lys Asn	
215 220 225	
aac cca gtt ctt att ggt gag cca ggt gtt ggt aag acc gca gtt gtt	835
Asn Pro Val Leu Ile Gly Glu Pro Gly Val Gly Lys Thr Ala Val Val	
230 235 240 245	
gaa ggt ctt gca cta gac att gtt aac ggc aag gtt cca gag acc ctc	883
Glu Gly Leu Ala Leu Asp Ile Val Asn Gly Lys Val Pro Glu Thr Leu	
250 255 260	
aag gac aag cag gtt tac tcc ctt gac tta ggt tcc ctg gtt gca ggt	931
Lys Asp Lys Gln Val Tyr Ser Leu Asp Leu Gly Ser Leu Val Ala Gly	
265 270 275	
tcc cgt tac cgc ggt gac ttc gaa gag cga ctg aag aag gtc ctc aag	979
Ser Arg Tyr Arg Gly Asp Phe Glu Glu Arg Leu Lys Lys Val Leu Lys	
280 285 290	
gag att aac cag cgc ggc gac atc atc ctg ttt atc gat gag atc cac	1027
Glu Ile Asn Gln Arg Gly Asp Ile Ile Leu Phe Ile Asp Glu Ile His	
295 300 305	
acc ctc gtg ggt gca ggt gca gca gaa ggc gca atc gac gct gcc tcc	1075
Thr Leu Val Gly Ala Gly Ala Ala Glu Gly Ala Ile Asp Ala Ala Ser	
310 315 320 325	
ctg ctt aag cca aag ctt gcc cgc ggt gaa ctg cag acc att ggt gca	1123
Leu Leu Lys Pro Lys Leu Ala Arg Gly Glu Leu Gln Thr Ile Gly Ala	
330 335 340	
acc acc ctg gat gag tac cgt aag cac att gaa aag gac gca gct ctt	1171
Thr Thr Leu Asp Glu Tyr Arg Lys His Ile Glu Lys Asp Ala Ala Leu	
345 350 355	
gag cgt cgt ttc cag cca gtg cag gtt cca gag cct tcg gtt gat ctc	1219
Glu Arg Arg Phe Gln Pro Val Gln Val Pro Glu Pro Ser Val Asp Leu	
360 365 370	
acc gtt gag atc ttg aag ggt ctg cgc gac cgc tac gaa gct cac cac	1267
Thr Val Glu Ile Leu Lys Gly Leu Arg Asp Arg Tyr Glu Ala His His	
375 380 385	
cgc gta tcc atc acc gat ggt gct ctt act gca gca gct cag ctt gct	1315
Arg Val Ser Ile Thr Asp Gly Ala Leu Thr Ala Ala Ala Gln Leu Ala	
390 395 400 405	

gat cgc tac atc aac gac cgc ttc ttg cca gat aag gcc gtt gac ctc	1363
Asp Arg Tyr Ile Asn Asp Arg Phe Leu Pro Asp Lys Ala Val Asp Leu	
410 415 420	
atc gat gag gct ggc gcc cgc atg cgc atc aag cgc atg acc gca cct	1411
Ile Asp Glu Ala Gly Ala Arg Met Arg Ile Lys Arg Met Thr Ala Pro	
425 430 435	
tcc tcc ctc cgc gag gtt gat gag cgt atc gct gat gtt cgc cgt gag	1459
Ser Ser Leu Arg Glu Val Asp Glu Arg Ile Ala Asp Val Arg Arg Glu	
440 445 450	
aag gaa gca gcg atc gat gct cag gac ttt gag aag gca gca ggt ctt	1507
Lys Glu Ala Ala Ile Asp Ala Gln Asp Phe Glu Lys Ala Ala Gly Leu	
455 460 465	
cgc gat aag gag cgc aag ctc ggc gaa gag cgt tca gag aag gaa aag	1555
Arg Asp Lys Glu Arg Lys Leu Gly Glu Glu Arg Ser Glu Lys Glu Lys	
470 475 480 485	
cag tgg cgc tcc ggc gac ctc gag gac atc gct gag gtt ggc gaa gag	1603
Gln Trp Arg Ser Gly Asp Leu Glu Asp Ile Ala Glu Val Gly Glu Glu	
490 495 500	
cag atc gca gaa gta ctg gcc aac tgg act ggt att cct gtc ttc aag	1651
Gln Ile Ala Glu Val Leu Ala Asn Trp Thr Gly Ile Pro Val Phe Lys	
505 510 515	
ctc acc gaa gct gaa tct tca cgc ctg ctc aac atg gaa gaa gag ttg	1699
Leu Thr Glu Ala Glu Ser Ser Arg Leu Leu Asn Met Glu Glu Glu Leu	
520 525 530	
cac aag cgc atc atc gga cag gat gaa gct gtc aag gct gtc tcc cgt	1747
His Lys Arg Ile Ile Gly Gln Asp Glu Ala Val Lys Ala Val Ser Arg	
535 540 545	
gcg atc cgt cgt acc cgt gca ggt ctg aag gat cct aag cgt cct tcc	1795
Ala Ile Arg Arg Thr Arg Ala Gly Leu Lys Asp Pro Lys Arg Pro Ser	
550 555 560 565	
ggc tcc ttc atc ttc gct ggt cca tcc ggc gtt ggt aag acc gag ctg	1843
Gly Ser Phe Ile Phe Ala Gly Pro Ser Gly Val Gly Lys Thr Glu Leu	
570 575 580	
tcc aag gct ctc gca gga ttc ctc ttc ggt gac gat gat tcc ctc atc	1891
Ser Lys Ala Leu Ala Gly Phe Leu Phe Gly Asp Asp Asp Ser Leu Ile	
585 590 595	
caa atc gac atg ggt gag ttc cac gac cgc ttc acc gcg tcc cga ctt	1939
Gln Ile Asp Met Gly Glu Phe His Asp Arg Phe Thr Ala Ser Arg Leu	
600 605 610	
ttc ggt gcc cct ccg gga tac gtt ggc tac gaa gaa ggt ggc cag ctg	1987
Phe Gly Ala Pro Pro Gly Tyr Val Gly Tyr Glu Glu Gly Gly Gln Leu	
615 620 625	
acc gag aag gtt cgc cgt aag cca ttc tcc gtt gtg ctt ttc gac gaa	2035
Thr Glu Lys Val Arg Arg Lys Pro Phe Ser Val Val Leu Phe Asp Glu	
630 635 640 645	
atc gag aag gcc cac aag gag atc tac aac acc ttg ctg cag gtg ttg	2083

Ile Glu Lys Ala His Lys Glu Ile Tyr Asn Thr Leu Leu Gln Val Leu	
650 655 660	
gaa gat ggt cgc ctt acc gat ggt cag gga cgc atc gtg gac ttc aag	2131
Glu Asp Gly Arg Leu Thr Asp Gly Gln Gly Arg Ile Val Asp Phe Lys	
665 670 675	
aac acc gtc ctg atc ttc acc tcc aac ctg ggc acc gct gac atc tcc	2179
Asn Thr Val Leu Ile Phe Thr Ser Asn Leu Gly Thr Ala Asp Ile Ser	
680 685 690	
aag gct gtt ggc ctg ggc ttc tcc gga tcc tcc gag act gac agc gat	2227
Lys Ala Val Gly Leu Gly Phe Ser Gly Ser Ser Glu Thr Asp Ser Asp	
695 700 705	
gct cag tac gac cgc atg aag aac aag gtc cac gac gag ctg aag aag	2275
Ala Gln Tyr Asp Arg Met Lys Asn Lys Val His Asp Glu Leu Lys Lys	
710 715 720 725	
cac ttc cgc cct gag ttc ctg aac cgt att gat gag atc gtg gtc ttc	2323
His Phe Arg Pro Glu Phe Leu Asn Arg Ile Asp Glu Ile Val Val Phe	
730 735 740	
cac cag ctc acc aag gat cag atc gtt cag atg gtc gac ctt ctt atc	2371
His Gln Leu Thr Lys Asp Gln Ile Val Gln Met Val Asp Leu Leu Ile	
745 750 755	
ggt cgc gtt tcc aac gca ctg gct gag aag gac atg agc atc gaa ctg	2419
Gly Arg Val Ser Asn Ala Leu Ala Glu Lys Asp Met Ser Ile Glu Leu	
760 765 770	
act gag aag gcc aag gac ctc ctg gct aac cga ggc ttc gat cca gtt	2467
Thr Glu Lys Ala Lys Asp Leu Leu Ala Asn Arg Gly Phe Asp Pro Val	
775 780 785	
ctg ggt gca cga cca ttg cgt cgc acc atc cag cgc gaa att gaa gac	2515
Leu Gly Ala Arg Pro Leu Arg Arg Thr Ile Gln Arg Glu Ile Glu Asp	
790 795 800 805	
cag atg tcc gag aag atc ctc ttc ggt gaa atc ggc gca ggc gag atc	2563
Gln Met Ser Glu Lys Ile Leu Phe Gly Glu Ile Gly Ala Gly Glu Ile	
810 815 820	
gtc acc gtt gac gtc gaa ggc tgg gac ggc gag tcc aag gac acc gac	2611
Val Thr Val Asp Val Glu Gly Trp Asp Gly Glu Ser Lys Asp Thr Asp	
825 830 835	
cgt gcg aag ttc acc ttc aca cca cgt cca aag cca atg cca gaa ggt	2659
Arg Ala Lys Phe Thr Phe Thr Pro Arg Pro Lys Pro Met Pro Glu Gly	
840 845 850	
aag ttc tct gag atc tct gtc gag gct gcg gaa gca att caa gat gta	2707
Lys Phe Ser Glu Ile Ser Val Glu Ala Ala Glu Ala Ile Gln Asp Val	
855 860 865	
gat tct gca gct gac ggc gat gtc cca gaa acc gat tca ctt tcc gac	2755
Asp Ser Ala Ala Asp Gly Asp Val Pro Glu Thr Asp Ser Leu Ser Asp	
870 875 880 885	
att gac ctt gaa acc ctt gaa aag ttt gag gaa gat gta gaa aac ggc	2803
Ile Asp Leu Glu Thr Leu Glu Lys Phe Glu Glu Asp Val Glu Asn Gly	

890 895 900  
 acc gac att gat cag gtg tcc ggt gac tac tac ggc acc gat gat cag 2851  
 Thr Asp Ile Asp Gln Val Ser Gly Asp Tyr Tyr Gly Thr Asp Asp Gln  
 905 910 915

gga ggc act gct cca agc aag gag tagcaacott ttgaaaaagg gcc 2898  
 Gly Gly Thr Ala Pro Ser Lys Glu  
 920 925

<210> 50

<211> 925

<212> PRT

<213> Corynebacterium glutamicum

<400> 50

Met Phe Glu Arg Phe Thr Asp Arg Ala Arg Arg Val Ile Val Leu Ala  
 1 5 10 15

Gln Glu Glu Ala Arg Met Leu Asn His Asn Tyr Ile Gly Thr Glu His  
 20 25 30

Ile Leu Leu Gly Leu Ile His Glu Gly Glu Gly Val Ala Ala Lys Ala  
 35 40 45

Leu Glu Ser Met Gly Ile Ser Leu Asp Ala Val Arg Gln Glu Val Glu  
 50 55 60

Glu Ile Ile Gly Gln Gly Ser Gln Pro Thr Thr Gly His Ile Pro Phe  
 65 70 75 80

Thr Pro Arg Ala Lys Lys Val Leu Glu Leu Ser Leu Arg Glu Gly Leu  
 85 90 95

Gln Met Gly His Lys Tyr Ile Gly Thr Glu Phe Leu Leu Leu Gly Leu  
 100 105 110

Ile Arg Glu Gly Glu Gly Val Ala Ala Gln Val Leu Val Lys Leu Gly  
 115 120 125

Ala Asp Leu Pro Arg Val Arg Gln Gln Val Ile Gln Leu Leu Ser Gly  
 130 135 140

Tyr Glu Gly Gly Gln Gly Gly Ser Pro Glu Gly Gly Gln Gly Ala Pro  
 145 150 155 160

Thr Gly Gly Asp Ala Val Gly Ala Gly Ala Ala Pro Gly Gly Arg Pro  
 165 170 175

Ser Ser Gly Ser Pro Gly Glu Arg Ser Thr Ser Leu Val Leu Asp Gln  
 180 185 190

Phe Gly Arg Asn Leu Thr Gln Ala Ala Lys Asp Gly Lys Leu Asp Pro  
 195 200 205

Val Val Gly Arg Asp Lys Glu Ile Glu Arg Ile Met Gln Val Leu Ser  
 210 215 220

Arg Arg Thr Lys Asn Asn Pro Val Leu Ile Gly Glu Pro Gly Val Gly  
 225 230 235 240

70

Pro Lys Arg Pro Ser Gly Ser Phe Ile Phe Ala Gly Pro Ser Gly Val  
 565 570 575  
 Gly Lys Thr Glu Leu Ser Lys Ala Leu Ala Gly Phe Leu Phe Gly Asp  
 580 585 590  
 Asp Asp Ser Leu Ile Gln Ile Asp Met Gly Glu Phe His Asp Arg Phe  
 595 600 605  
 Thr Ala Ser Arg Leu Phe Gly Ala Pro Pro Gly Tyr Val Gly Tyr Glu  
 610 615 620  
 Glu Gly Gly Gln Leu Thr Glu Lys Val Arg Arg Lys Pro Phe Ser Val  
 625 630 635 640  
 Val Leu Phe Asp Glu Ile Glu Lys Ala His Lys Glu Ile Tyr Asn Thr  
 645 650 655  
 Leu Leu Gln Val Leu Glu Asp Gly Arg Leu Thr Asp Gly Gln Gly Arg  
 660 665 670  
 Ile Val Asp Phe Lys Asn Thr Val Leu Ile Phe Thr Ser Asn Leu Gly  
 675 680 685  
 Thr Ala Asp Ile Ser Lys Ala Val Gly Leu Gly Phe Ser Gly Ser Ser  
 690 695 700  
 Glu Thr Asp Ser Asp Ala Gln Tyr Asp Arg Met Lys Asn Lys Val His  
 705 710 715 720  
 Asp Glu Leu Lys Lys His Phe Arg Pro Glu Phe Leu Asn Arg Ile Asp  
 725 730 735  
 Glu Ile Val Val Phe His Gln Leu Thr Lys Asp Gln Ile Val Gln Met  
 740 745 750  
 Val Asp Leu Leu Ile Gly Arg Val Ser Asn Ala Leu Ala Glu Lys Asp  
 755 760 765  
 Met Ser Ile Glu Leu Thr Glu Lys Ala Lys Asp Leu Leu Ala Asn Arg  
 770 775 780  
 Gly Phe Asp Pro Val Leu Gly Ala Arg Pro Leu Arg Arg Thr Ile Gln  
 785 790 795 800  
 Arg Glu Ile Glu Asp Gln Met Ser Glu Lys Ile Leu Phe Gly Glu Ile  
 805 810 815  
 Gly Ala Gly Glu Ile Val Thr Val Asp Val Glu Gly Trp Asp Gly Glu  
 820 825 830  
 Ser Lys Asp Thr Asp Arg Ala Lys Phe Thr Phe Thr Pro Arg Pro Lys  
 835 840 845  
 Pro Met Pro Glu Gly Lys Phe Ser Glu Ile Ser Val Glu Ala Ala Glu  
 850 855 860  
 Ala Ile Gln Asp Val Asp Ser Ala Ala Asp Gly Asp Val Pro Glu Thr  
 865 870 875 880  
 Asp Ser Leu Ser Asp Ile Asp Leu Glu Thr Leu Glu Lys Phe Glu Glu

885										890					895				
Asp	Val	Glu	Asn	Gly	Thr	Asp	Ile	Asp	Gln	Val	Ser	Gly	Asp	Tyr	Tyr				
			900					905					910						
Gly	Thr	Asp	Asp	Gln	Gly	Gly	Thr	Ala	Pro	Ser	Lys	Glu							
			915				920					925							
 <210> 51																			
<211> 1104																			
<212> DNA																			
<213> Corynebacterium glutamicum																			
 <220>																			
<221> CDS																			
<222> (101)..(1081)																			
<223> FRXA02470																			
 <400> 51																			
ctaattccaa gccgagctga aaaagtctgg aagttttgcc caataagggc gttaaagtgg 60																			
gtgaaagcga atttagaaat aaagaattaa ggggagagac atg ttc gag agg ttt 115																			
Met Phe Glu Arg Phe 1 5																			
acc gat cgt gca cgc cgc gtg att gtg ctc gcg cag gaa gag gcg cgc 163																			
Thr Asp Arg Ala Arg Arg Val Ile Val Leu Ala Gln Glu Glu Ala Arg 10 15 20																			
atg ctc aac cac aat tac atc ggc acg gag cac att ctc ctc ggc ctc 211																			
Met Leu Asn His Asn Tyr Ile Gly Thr Glu His Ile Leu Leu Gly Leu 25 30 35																			
att cac gag ggc gag ggc gtt gca gcc aag gct ttg gaa tcc atg gga 259																			
Ile His Glu Gly Glu Gly Val Ala Ala Lys Ala Leu Glu Ser Met Gly 40 45 50																			
att tcc ctg gac gcc gtc cgc cag gaa gtc gaa gag att atc ggc cag 307																			
Ile Ser Leu Asp Ala Val Arg Gln Glu Val Glu Glu Ile Ile Gly Gln 55 60 65																			
ggc tcc cag ccc acc acc ggc cac att cct ttt act cca cgt gcc aag 355																			
Gly Ser Gln Pro Thr Thr Gly His Ile Pro Phe Thr Pro Arg Ala Lys 70 75 80 85																			
aag gtc ctg gag ctc agc ctc cgc gaa ggc cta caa atg gga cac aag 403																			
Lys Val Leu Glu Leu Ser Leu Arg Glu Gly Leu Gln Met Gly His Lys 90 95 100																			
tac atc ggt act gag ttc ctg ctt ctc ggt ttg atc cgt gag ggc gag 451																			
Tyr Ile Gly Thr Glu Phe Leu Leu Leu Gly Leu Ile Arg Glu Gly Glu 105 110 115																			
ggc gtt gct gcc cag gtc ctg gtc aag ctt ggt gct gat ctg cca cgc 499																			
Gly Val Ala Ala Gln Val Leu Val Lys Leu Gly Ala Asp Leu Pro Arg 120 125 130																			
gtg cgt cag caa gtt att cag ctt ctc tcc ggc tac gaa ggt ggc cag 547																			
Val Arg Gln Gln Val Ile Gln Leu Leu Ser Gly Tyr Glu Gly Gly Gln 135 140 145																			



ggc gga tcc cca gag ggc ggc cag ggc gcc cct act ggc ggt gac gct 595  
 Gly Gly Ser Pro Glu Gly Gly Gln Gly Ala Pro Thr Gly Gly Asp Ala  
 150 155 160 165  
 gtt ggt gca gga gct gct cct ggc ggt cgt cca tct tcg ggc agc cca 643  
 Val Gly Ala Gly Ala Ala Pro Gly Gly Arg Pro Ser Ser Gly Ser Pro  
 170 175 180  
 ggc gag cgt tct acc tct ttg gtc ctt gac cag ttc ggc cgc aac ctc 691  
 Gly Glu Arg Ser Thr Ser Leu Val Leu Asp Gln Phe Gly Arg Asn Leu  
 185 190 195  
 acc cag gct gca aag gac ggc aag ctg gat cca gtt gtt ggt cgc gat 739  
 Thr Gln Ala Ala Lys Asp Gly Lys Leu Asp Pro Val Val Gly Arg Asp  
 200 205 210  
 aag gaa atc gag cgc atc atg cag gtg ctc tcc cgt cgt acc aag aac 787  
 Lys Glu Ile Glu Arg Ile Met Gln Val Leu Ser Arg Arg Thr Lys Asn  
 215 220 225  
 aac cca gtt ctt att ggt gag cca ggt gtt ggt aag acc gca gtt gtt 835  
 Asn Pro Val Leu Ile Gly Glu Pro Gly Val Gly Lys Thr Ala Val Val  
 230 235 240 245  
 gaa ggt ctt gca cta gac att gtt aac ggc aag gtt cca gag acc ctc 883  
 Glu Gly Leu Ala Leu Asp Ile Val Asn Gly Lys Val Pro Glu Thr Leu  
 250 255 260  
 aag gac aag cag gtt tac tcc ctt gac tta ggt tcc ctg gtt gca ggt 931  
 Lys Asp Lys Gln Val Tyr Ser Leu Asp Leu Gly Ser Leu Val Ala Gly  
 265 270 275  
 tcc cgt tac cgc ggt gac ttc gaa gag cga ctg aag aag gtc ctc aag 979  
 Ser Arg Tyr Arg Gly Asp Phe Glu Glu Arg Leu Lys Lys Val Leu Lys  
 280 285 290  
 gag att aac cag cgc ggc gac atc atc ctg ttt atc gat gag atc cac 1027  
 Glu Ile Asn Gln Arg Gly Asp Ile Ile Leu Phe Ile Asp Glu Ile His  
 295 300 305  
 acc ctc gtg ggt gca ggt gca gca cga agg cgc aat cga cgc tgc ctc 1075  
 Thr Leu Val Gly Ala Gly Ala Ala Arg Arg Arg Asn Arg Arg Cys Leu  
 310 315 320 325  
 cct gct taagccaaag cttgcccgcg gtg 1104  
 Pro Ala

&lt;210&gt; 52

&lt;211&gt; 327

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 52

Met Phe Glu Arg Phe Thr Asp Arg Ala Arg Arg Val Ile Val Leu Ala  
 1 5 10 15

Gln Glu Glu Ala Arg Met Leu Asn His Asn Tyr Ile Gly Thr Glu His  
 20 25 30

Ile Leu Leu Gly Leu Ile His Glu Gly Glu Gly Val Ala Ala Lys Ala  
 35 40 45  
 Leu Glu Ser Met Gly Ile Ser Leu Asp Ala Val Arg Gln Glu Val Glu  
 50 55 60  
 Glu Ile Ile Gly Gln Gly Ser Gln Pro Thr Thr Gly His Ile Pro Phe  
 65 70 75 80  
 Thr Pro Arg Ala Lys Lys Val Leu Glu Leu Ser Leu Arg Glu Gly Leu  
 85 90 95  
 Gln Met Gly His Lys Tyr Ile Gly Thr Glu Phe Leu Leu Leu Gly Leu  
 100 105 110  
 Ile Arg Glu Gly Glu Gly Val Ala Ala Gln Val Leu Val Lys Leu Gly  
 115 120 125  
 Ala Asp Leu Pro Arg Val Arg Gln Gln Val Ile Gln Leu Leu Ser Gly  
 130 135 140  
 Tyr Glu Gly Gly Gln Gly Gly Ser Pro Glu Gly Gly Gln Gly Ala Pro  
 145 150 155 160  
 Thr Gly Gly Asp Ala Val Gly Ala Gly Ala Ala Pro Gly Gly Arg Pro  
 165 170 175  
 Ser Ser Gly Ser Pro Gly Glu Arg Ser Thr Ser Leu Val Leu Asp Gln  
 180 185 190  
 Phe Gly Arg Asn Leu Thr Gln Ala Ala Lys Asp Gly Lys Leu Asp Pro  
 195 200 205  
 Val Val Gly Arg Asp Lys Glu Ile Glu Arg Ile Met Gln Val Leu Ser  
 210 215 220  
 Arg Arg Thr Lys Asn Asn Pro Val Leu Ile Gly Glu Pro Gly Val Gly  
 225 230 235 240  
 Lys Thr Ala Val Val Glu Gly Leu Ala Leu Asp Ile Val Asn Gly Lys  
 245 250 255  
 Val Pro Glu Thr Leu Lys Asp Lys Gln Val Tyr Ser Leu Asp Leu Gly  
 260 265 270  
 Ser Leu Val Ala Gly Ser Arg Tyr Arg Gly Asp Phe Glu Glu Arg Leu  
 275 280 285  
 Lys Lys Val Leu Lys Glu Ile Asn Gln Arg Gly Asp Ile Ile Leu Phe  
 290 295 300  
 Ile Asp Glu Ile His Thr Leu Val Gly Ala Gly Ala Ala Arg Arg Arg  
 305 310 315 320  
 Asn Arg Arg Cys Leu Pro Ala  
 325

&lt;210&gt; 53

&lt;211&gt; 1956

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (101)..(1933)

&lt;223&gt; FRXA02471

&lt;400&gt; 53

```

ggtgacttcg aagagcgact gaagaagggtc ctcaaggaga ttaaccagcg cggcgacatc 60

atcctgttta tcgatgagat ccacaccctc gtgggtgcag  gtg  cag  cac  gaa  ggc    115
                               Val  Gln  His  Glu  Gly
                               1      5

gca atc gac gct gcc tcc ctg ctt aag cca aag ctt gcc cgc ggt gaa    163
Ala Ile Asp Ala Ala Ser Leu Leu Lys Pro Lys Leu Ala Arg Gly Glu
                      10                      15                      20

ctg cag acc att ggt gca acc acc ctg gat gag tac cgt aag cac att    211
Leu Gln Thr Ile Gly Ala Thr Thr Leu Asp Glu Tyr Arg Lys His Ile
                      25                      30                      35

gaa aag gac gca gct ctt gag cgt cgt ttc cag cca gtg cag gtt cca    259
Glu Lys Asp Ala Ala Leu Glu Arg Arg Phe Gln Pro Val Gln Val Pro
                      40                      45                      50

gag cct tcg gtt gat ctc acc gtt gag atc ttg aag ggt ctg cgc gac    307
Glu Pro Ser Val Asp Leu Thr Val Glu Ile Leu Lys Gly Leu Arg Asp
                      55                      60                      65

cgc tac gaa gct cac cac cgc gta tcc atc acc gat ggt gct ctt act    355
Arg Tyr Glu Ala His His Arg Val Ser Ile Thr Asp Gly Ala Leu Thr
                      70                      75                      80                      85

gca gca gct cag ctt gct gat cgc tac atc aac gac cgc ttc ttg cca    403
Ala Ala Ala Gln Leu Ala Asp Arg Tyr Ile Asn Asp Arg Phe Leu Pro
                      90                      95                      100

gat aag gcc gtt gac ctc atc gat gag gct ggc gcc cgc atg cgc atc    451
Asp Lys Ala Val Asp Leu Ile Asp Glu Ala Gly Ala Arg Met Arg Ile
                      105                      110                      115

aag cgc atg acc gca cct tcc tcc ctc cgc gag gtt gat gag cgt atc    499
Lys Arg Met Thr Ala Pro Ser Ser Leu Arg Glu Val Asp Glu Arg Ile
                      120                      125                      130

gct gat gtt cgc cgt gag aag gaa gca gcg atc gat gct cag gac ttt    547
Ala Asp Val Arg Arg Glu Lys Glu Ala Ala Ile Asp Ala Gln Asp Phe
                      135                      140                      145

gag aag gca gca ggt ctt cgc gat aag gag cgc aag ctc ggc gaa gag    595
Glu Lys Ala Ala Gly Leu Arg Asp Lys Glu Arg Lys Leu Gly Glu Glu
                      150                      155                      160                      165

cgt tca gag aag gaa aag cag tgg cgc tcc ggc gac ctc gag gac atc    643
Arg Ser Glu Lys Glu Lys Gln Trp Arg Ser Gly Asp Leu Glu Asp Ile
                      170                      175                      180

gct gag gtt ggc gaa gag cag atc gca gaa gta ctg gcc aac tgg act    691
Ala Glu Val Gly Glu Glu Gln Ile Ala Glu Val Leu Ala Asn Trp Thr

```

185	190	195	
ggt att cct gtc ttc aag ctc acc gaa gct gaa tct tca cgc ctg ctc Gly Ile Pro Val Phe Lys Leu Thr Glu Ala Glu Ser Ser Arg Leu Leu 200 205 210			739
aac atg gaa gaa gag ttg cac aag cgc atc atc gga cag gat gaa gct Asn Met Glu Glu Glu Leu His Lys Arg Ile Ile Gly Gln Asp Glu Ala 215 220 225			787
gtc aag gct gtc tcc cgt gcg atc cgt cgt acc cgt gca ggt ctg aag Val Lys Ala Val Ser Arg Ala Ile Arg Arg Thr Arg Ala Gly Leu Lys 230 235 240 245			835
gat cct aag cgt cct tcc ggc tcc ttc atc ttc gct ggt cca tcc ggc Asp Pro Lys Arg Pro Ser Gly Ser Phe Ile Phe Ala Gly Pro Ser Gly 250 255 260			883
gtt ggt aag acc gag ctg tcc aag gct ctc gca gga ttc ctc ttc ggt Val Gly Lys Thr Glu Leu Ser Lys Ala Leu Ala Gly Phe Leu Phe Gly 265 270 275			931
gac gat gat tcc ctc atc caa atc gac atg ggt gag ttc cac gac cgc Asp Asp Asp Ser Leu Ile Gln Ile Asp Met Gly Glu Phe His Asp Arg 280 285 290			979
ttc acc gcg tcc cga ctt ttc ggt gcc cct ccg gga tac gtt ggc tac Phe Thr Ala Ser Arg Leu Phe Gly Ala Pro Pro Gly Tyr Val Gly Tyr 295 300 305			1027
gaa gaa ggt ggc cag ctg acc gag aag gtt cgc cgt aag cca ttc tcc Glu Glu Gly Gly Gln Leu Thr Glu Lys Val Arg Arg Lys Pro Phe Ser 310 315 320 325			1075
gtt gtg ctt ttc gac gaa atc gag aag gcc cac aag gag atc tac aac Val Val Leu Phe Asp Glu Ile Glu Lys Ala His Lys Glu Ile Tyr Asn 330 335 340			1123
acc ttg ctg cag gtg ttg gaa gat ggt cgc ctt acc gat ggt cag gga Thr Leu Leu Gln Val Leu Glu Asp Gly Arg Leu Thr Asp Gly Gln Gly 345 350 355			1171
cgc atc gtg gac ttc aag aac acc gtc ctg atc ttc acc tcc aac ctg Arg Ile Val Asp Phe Lys Asn Thr Val Leu Ile Phe Thr Ser Asn Leu 360 365 370			1219
ggc acc gct gac atc tcc aag gct gtt ggc ctg ggc ttc tcc gga tcc Gly Thr Ala Asp Ile Ser Lys Ala Val Gly Leu Gly Phe Ser Gly Ser 375 380 385			1267
tcc gag act gac agc gat gct cag tac gac cgc atg aag aac aag gtc Ser Glu Thr Asp Ser Asp Ala Gln Tyr Asp Arg Met Lys Asn Lys Val 390 395 400 405			1315
cac gac gag ctg aag aag cac ttc cgc cct gag ttc ctg aac cgt att His Asp Glu Leu Lys Lys His Phe Arg Pro Glu Phe Leu Asn Arg Ile 410 415 420			1363
gat gag atc gtg gtc ttc cac cag ctc acc aag gat cag atc gtt cag Asp Glu Ile Val Val Phe His Gln Leu Thr Lys Asp Gln Ile Val Gln 425 430 435			1411

atg gtc gac ctt ctt atc ggt cgc gtt tcc aac gca ctg gct gag aag 1459  
 Met Val Asp Leu Leu Ile Gly Arg Val Ser Asn Ala Leu Ala Glu Lys  
 440 445 450  
 gac atg agc atc gaa ctg act gag aag gcc aag gac ctc ctg gct aac 1507  
 Asp Met Ser Ile Glu Leu Thr Glu Lys Ala Lys Asp Leu Leu Ala Asn  
 455 460 465  
 cga ggc ttc gat cca gtt ctg ggt gca cga cca ttg cgt cgc acc atc 1555  
 Arg Gly Phe Asp Pro Val Leu Gly Ala Arg Pro Leu Arg Arg Thr Ile  
 470 475 480 485  
 cag cgc gaa att gaa gac cag atg tcc gag aag atc ctc ttc ggt gaa 1603  
 Gln Arg Glu Ile Glu Asp Gln Met Ser Glu Lys Ile Leu Phe Gly Glu  
 490 495 500  
 atc ggc gca ggc gag atc gtc acc gtt gac gtc gaa ggc tgg gac ggc 1651  
 Ile Gly Ala Gly Glu Ile Val Thr Val Asp Val Glu Gly Trp Asp Gly  
 505 510 515  
 gag tcc aag gac acc gac cgt gcg aag ttc acc ttc aca cca cgt cca 1699  
 Glu Ser Lys Asp Thr Asp Arg Ala Lys Phe Thr Phe Thr Pro Arg Pro  
 520 525 530  
 aag cca atg cca gaa ggt aag ttc tct gag atc tct gtc gag gct gcg 1747  
 Lys Pro Met Pro Glu Gly Lys Phe Ser Glu Ile Ser Val Glu Ala Ala  
 535 540 545  
 gaa gca att caa gat gta gat tct gca gct gac ggc gat gtc cca gaa 1795  
 Glu Ala Ile Gln Asp Val Asp Ser Ala Ala Asp Gly Asp Val Pro Glu  
 550 555 560 565  
 acc gat tca ctt tcc gac att gac ctt gaa acc ctt gaa aag ttt gag 1843  
 Thr Asp Ser Leu Ser Asp Ile Asp Leu Glu Thr Leu Glu Lys Phe Glu  
 570 575 580  
 gaa gat gta gaa aac ggc acc gac att gat cag gtg tcc ggt gac tac 1891  
 Glu Asp Val Glu Asn Gly Thr Asp Ile Asp Gln Val Ser Gly Asp Tyr  
 585 590 595  
 tac ggc acc gat gat cag gga ggc act gct cca agc aag gag 1933  
 Tyr Gly Thr Asp Asp Gln Gly Gly Thr Ala Pro Ser Lys Glu  
 600 605 610  
 tagcaacctt ttgaaaaagg gcc 1956

<210> 54  
 <211> 611  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 54  
 Val Gln His Glu Gly Ala Ile Asp Ala Ala Ser Leu Leu Lys Pro Lys  
 1 5 10 15  
 Leu Ala Arg Gly Glu Leu Gln Thr Ile Gly Ala Thr Thr Leu Asp Glu  
 20 25 30  
 Tyr Arg Lys His Ile Glu Lys Asp Ala Ala Leu Glu Arg Arg Phe Gln

35					40					45					
Pro	Val	Gln	Val	Pro	Glu	Pro	Ser	Val	Asp	Leu	Thr	Val	Glu	Ile	Leu
50						55					60				
Lys	Gly	Leu	Arg	Asp	Arg	Tyr	Glu	Ala	His	His	Arg	Val	Ser	Ile	Thr
65					70					75					80
Asp	Gly	Ala	Leu	Thr	Ala	Ala	Ala	Gln	Leu	Ala	Asp	Arg	Tyr	Ile	Asn
				85					90					95	
Asp	Arg	Phe	Leu	Pro	Asp	Lys	Ala	Val	Asp	Leu	Ile	Asp	Glu	Ala	Gly
			100					105					110		
Ala	Arg	Met	Arg	Ile	Lys	Arg	Met	Thr	Ala	Pro	Ser	Ser	Leu	Arg	Glu
		115					120					125			
Val	Asp	Glu	Arg	Ile	Ala	Asp	Val	Arg	Arg	Glu	Lys	Glu	Ala	Ala	Ile
	130					135					140				
Asp	Ala	Gln	Asp	Phe	Glu	Lys	Ala	Ala	Gly	Leu	Arg	Asp	Lys	Glu	Arg
145					150					155					160
Lys	Leu	Gly	Glu	Glu	Arg	Ser	Glu	Lys	Glu	Lys	Gln	Trp	Arg	Ser	Gly
			165						170					175	
Asp	Leu	Glu	Asp	Ile	Ala	Glu	Val	Gly	Glu	Glu	Gln	Ile	Ala	Glu	Val
			180					185					190		
Leu	Ala	Asn	Trp	Thr	Gly	Ile	Pro	Val	Phe	Lys	Leu	Thr	Glu	Ala	Glu
		195					200					205			
Ser	Ser	Arg	Leu	Leu	Asn	Met	Glu	Glu	Glu	Leu	His	Lys	Arg	Ile	Ile
	210					215					220				
Gly	Gln	Asp	Glu	Ala	Val	Lys	Ala	Val	Ser	Arg	Ala	Ile	Arg	Arg	Thr
225					230					235					240
Arg	Ala	Gly	Leu	Lys	Asp	Pro	Lys	Arg	Pro	Ser	Gly	Ser	Phe	Ile	Phe
			245						250					255	
Ala	Gly	Pro	Ser	Gly	Val	Gly	Lys	Thr	Glu	Leu	Ser	Lys	Ala	Leu	Ala
			260					265					270		
Gly	Phe	Leu	Phe	Gly	Asp	Asp	Asp	Ser	Leu	Ile	Gln	Ile	Asp	Met	Gly
	275						280					285			
Glu	Phe	His	Asp	Arg	Phe	Thr	Ala	Ser	Arg	Leu	Phe	Gly	Ala	Pro	Pro
	290					295					300				
Gly	Tyr	Val	Gly	Tyr	Glu	Glu	Gly	Gly	Gln	Leu	Thr	Glu	Lys	Val	Arg
305					310					315					320
Arg	Lys	Pro	Phe	Ser	Val	Val	Leu	Phe	Asp	Glu	Ile	Glu	Lys	Ala	His
			325						330					335	
Lys	Glu	Ile	Tyr	Asn	Thr	Leu	Leu	Gln	Val	Leu	Glu	Asp	Gly	Arg	Leu
			340					345					350		
Thr	Asp	Gly	Gln	Gly	Arg	Ile	Val	Asp	Phe	Lys	Asn	Thr	Val	Leu	Ile
	355						360					365			

Phe Thr Ser Asn Leu Gly Thr Ala Asp Ile Ser Lys Ala Val Gly Leu  
 370 375 380  
 Gly Phe Ser Gly Ser Ser Glu Thr Asp Ser Asp Ala Gln Tyr Asp Arg  
 385 390 395 400  
 Met Lys Asn Lys Val His Asp Glu Leu Lys Lys His Phe Arg Pro Glu  
 405 410 415  
 Phe Leu Asn Arg Ile Asp Glu Ile Val Val Phe His Gln Leu Thr Lys  
 420 425 430  
 Asp Gln Ile Val Gln Met Val Asp Leu Leu Ile Gly Arg Val Ser Asn  
 435 440 445  
 Ala Leu Ala Glu Lys Asp Met Ser Ile Glu Leu Thr Glu Lys Ala Lys  
 450 455 460  
 Asp Leu Leu Ala Asn Arg Gly Phe Asp Pro Val Leu Gly Ala Arg Pro  
 465 470 475 480  
 Leu Arg Arg Thr Ile Gln Arg Glu Ile Glu Asp Gln Met Ser Glu Lys  
 485 490 495  
 Ile Leu Phe Gly Glu Ile Gly Ala Gly Glu Ile Val Thr Val Asp Val  
 500 505 510  
 Glu Gly Trp Asp Gly Glu Ser Lys Asp Thr Asp Arg Ala Lys Phe Thr  
 515 520 525  
 Phe Thr Pro Arg Pro Lys Pro Met Pro Glu Gly Lys Phe Ser Glu Ile  
 530 535 540  
 Ser Val Glu Ala Ala Glu Ala Ile Gln Asp Val Asp Ser Ala Ala Asp  
 545 550 555 560  
 Gly Asp Val Pro Glu Thr Asp Ser Leu Ser Asp Ile Asp Leu Glu Thr  
 565 570 575  
 Leu Glu Lys Phe Glu Glu Asp Val Glu Asn Gly Thr Asp Ile Asp Gln  
 580 585 590  
 Val Ser Gly Asp Tyr Tyr Gly Thr Asp Asp Gln Gly Gly Thr Ala Pro  
 595 600 605  
 Ser Lys Glu .  
 610

<210> 55  
 <211> 1446  
 <212> DNA  
 <213> *Corynebacterium glutamicum*

<220>  
 <221> CDS  
 <222> (101)..(1423)  
 <223> RXA02630

<400> 55

gcgggttcga aaatgtcgat gattaaacca cttaaagagct cacaggaagt gttcagacta 60

cttagagtga cgccccagcc acagggttca taatcaaatac atg aca aat caa ttc 115  
Met Thr Asn Gln Phe  
1 5

ccc aca aac aac ggt gag aac ccg gac cgt gca tcg gaa act cca tca 163  
Pro Thr Asn Asn Gly Glu Asn Pro Asp Arg Ala Ser Glu Thr Pro Ser  
10 15 20

gaa acc aac tcc ttc gaa cat gtg cgt agt tca tat ccg cag tgg ggt 211  
Glu Thr Asn Ser Phe Glu His Val Arg Ser Ser Tyr Pro Gln Trp Gly  
25 30 35

aac act gct tcc aat caa aac ccc tat cct ggt gcg ggc ttc ggc tct 259  
Asn Thr Ala Ser Asn Gln Asn Pro Tyr Pro Gly Ala Gly Phe Gly Ser  
40 45 50

gaa caa aac act caa caa gga aat gag caa caa gct cca gcc tgg acc 307  
Glu Gln Asn Thr Gln Gln Gly Asn Glu Gln Gln Ala Pro Ala Trp Thr  
55 60 65

agt tgg gat aat cag cct cta agc aca gat gta aag cca gct aaa gaa 355  
Ser Trp Asp Asn Gln Pro Leu Ser Thr Asp Val Lys Pro Ala Lys Glu  
70 75 80 85

aag cga aaa gtt ggc atc gga acg gca ctc gcg tta atg ctt gtt ggt 403  
Lys Arg Lys Val Gly Ile Gly Thr Ala Leu Ala Leu Met Leu Val Gly  
90 95 100

tct att gct acc ggt agc gtt gtg ggt gtt gca gca acc cag ctt ggt 451  
Ser Ile Ala Thr Gly Ser Val Val Gly Val Ala Ala Thr Gln Leu Gly  
105 110 115

tcg gac tct tca acc cca gtt aat gct ctt gag cag ccc agc gtg cag 499  
Ser Asp Ser Ser Thr Pro Val Asn Ala Leu Glu Gln Pro Ser Val Gln  
120 125 130

cgc acc act aat gct gaa cca ggt tca gcg gaa cag gtt gct gcc gca 547  
Arg Thr Thr Asn Ala Glu Pro Gly Ser Ala Glu Gln Val Ala Ala Ala  
135 140 145

gtt ttg oct tct gtc gtc tct att cag gcc att act agg acg tct gct 595  
Val Leu Pro Ser Val Val Ser Ile Gln Ala Ile Thr Arg Thr Ser Ala  
150 155 160 165

tct gag ggc tct gga tcc att att tcc tct gat ggt tac gtc atg acc 643  
Ser Glu Gly Ser Gly Ser Ile Ile Ser Ser Asp Gly Tyr Val Met Thr  
170 175 180

aat aat cac gtc gtg gca ggc att gaa caa tct ggt gtg tta gaa gta 691  
Asn Asn His Val Val Ala Gly Ile Glu Gln Ser Gly Val Leu Glu Val  
185 190 195

agt ttc tcc gat gga act aca gcg caa gct gat ttt att gct ggt gat 739  
Ser Phe Ser Asp Gly Thr Thr Ala Gln Ala Asp Phe Ile Ala Gly Asp  
200 205 210

cct tcc aca gat att gct gtg atc aag att agg gat gtg tcc aac ctt 787  
Pro Ser Thr Asp Ile Ala Val Ile Lys Ile Arg Asp Val Ser Asn Leu  
215 220 225



cca gtt atg agc ttt gga gat tct gac gca tta ggc gtt gga caa agt	835
Pro Val Met Ser Phe Gly Asp Ser Asp Ala Leu Gly Val Gly Gln Ser	
230 235 240 245	
gtg atg gct gtt ggt tct cca ctg ggt ctg agc tcc act gtg acc acc	883
Val Met Ala Val Gly Ser Pro Leu Gly Leu Ser Ser Thr Val Thr Thr	
250 255 260	
ggt att gtg tct gcc gtg aac cgt cct gtg cga gct tct ggt gat ggc	931
Gly Ile Val Ser Ala Val Asn Arg Pro Val Arg Ala Ser Gly Asp Gly	
265 270 275	
gga gag tct tcc ctc atc gat gct atc cag acc gat gct gcg atc aac	979
Gly Glu Ser Ser Leu Ile Asp Ala Ile Gln Thr Asp Ala Ala Ile Asn	
280 285 290	
cct ggt aac tct ggt ggt ccg ctg gtt gat atg gat ggc aac ctc att	1027
Pro Gly Asn Ser Gly Gly Pro Leu Val Asp Met Asp Gly Asn Leu Ile	
295 300 305	
ggc atg aat tct gta att gca tct att tct agc acc agc gat tcc gca	1075
Gly Met Asn Ser Val Ile Ala Ser Ile Ser Ser Thr Ser Asp Ser Ala	
310 315 320 325	
ggt tcc att ggt ctt ggt ttt tct atc cca tcc aac ttt gcc aag cgc	1123
Gly Ser Ile Gly Leu Gly Phe Ser Ile Pro Ser Asn Phe Ala Lys Arg	
330 335 340	
gtg gcc gat caa ttg atc agc acc ggc cag gta act cag ccg atg atc	1171
Val Ala Asp Gln Leu Ile Ser Thr Gly Gln Val Thr Gln Pro Met Ile	
345 350 355	
ggt gtg cag gtt ggc act gac aac tca gtg aca ggc gct gtg att gcc	1219
Gly Val Gln Val Gly Thr Asp Asn Ser Val Thr Gly Ala Val Ile Ala	
360 365 370	
agt gtt caa gat ggt gga ccg gcc gca gat gct gga ctt cag cca ggc	1267
Ser Val Gln Asp Gly Gly Pro Ala Ala Asp Ala Gly Leu Gln Pro Gly	
375 380 385	
gat atc gtg acc aag ctc aat gat cga gtt att gat agc cca gac tcc	1315
Asp Ile Val Thr Lys Leu Asn Asp Arg Val Ile Asp Ser Pro Asp Ser	
390 395 400 405	
ttg atc gct gct gtt cgt tct cat gat ttt ggc gaa acc gtc act tta	1363
Leu Ile Ala Ala Val Arg Ser His Asp Phe Gly Glu Thr Val Thr Leu	
410 415 420	
aca att aca cag cca gat acc tct cag agc cgg gag gta gag gtt act	1411
Thr Ile Thr Gln Pro Asp Thr Ser Gln Ser Arg Glu Val Glu Val Thr	
425 430 435	
ctg acg agt gag taggtttaaa agagttaatc tgc	1446
Leu Thr Ser Glu	
440	

<210> 56  
 <211> 441  
 <212> PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 56

```

Met Thr Asn Gln Phe Pro Thr Asn Asn Gly Glu Asn Pro Asp Arg Ala
 1           5           10           15

Ser Glu Thr Pro Ser Glu Thr Asn Ser Phe Glu His Val Arg Ser Ser
      20           25           30

Tyr Pro Gln Trp Gly Asn Thr Ala Ser Asn Gln Asn Pro Tyr Pro Gly
      35           40           45

Ala Gly Phe Gly Ser Glu Gln Asn Thr Gln Gln Gly Asn Glu Gln Gln
 50           55           60

Ala Pro Ala Trp Thr Ser Trp Asp Asn Gln Pro Leu Ser Thr Asp Val
 65           70           75           80

Lys Pro Ala Lys Glu Lys Arg Lys Val Gly Ile Gly Thr Ala Leu Ala
      85           90           95

Leu Met Leu Val Gly Ser Ile Ala Thr Gly Ser Val Val Gly Val Ala
      100           105           110

Ala Thr Gln Leu Gly Ser Asp Ser Ser Thr Pro Val Asn Ala Leu Glu
      115           120           125

Gln Pro Ser Val Gln Arg Thr Thr Asn Ala Glu Pro Gly Ser Ala Glu
      130           135           140

Gln Val Ala Ala Ala Val Leu Pro Ser Val Val Ser Ile Gln Ala Ile
      145           150           155           160

Thr Arg Thr Ser Ala Ser Glu Gly Ser Gly Ser Ile Ile Ser Ser Asp
      165           170           175

Gly Tyr Val Met Thr Asn Asn His Val Val Ala Gly Ile Glu Gln Ser
      180           185           190

Gly Val Leu Glu Val Ser Phe Ser Asp Gly Thr Thr Ala Gln Ala Asp
      195           200           205

Phe Ile Ala Gly Asp Pro Ser Thr Asp Ile Ala Val Ile Lys Ile Arg
      210           215           220

Asp Val Ser Asn Leu Pro Val Met Ser Phe Gly Asp Ser Asp Ala Leu
      225           230           235           240

Gly Val Gly Gln Ser Val Met Ala Val Gly Ser Pro Leu Gly Leu Ser
      245           250           255

Ser Thr Val Thr Thr Gly Ile Val Ser Ala Val Asn Arg Pro Val Arg
      260           265           270

Ala Ser Gly Asp Gly Gly Glu Ser Ser Leu Ile Asp Ala Ile Gln Thr
      275           280           285

Asp Ala Ala Ile Asn Pro Gly Asn Ser Gly Gly Pro Leu Val Asp Met
      290           295           300

Asp Gly Asn Leu Ile Gly Met Asn Ser Val Ile Ala Ser Ile Ser Ser

```

305	310	315	320
Thr Ser Asp Ser	Ala Gly Ser Ile Gly Leu Gly Phe Ser Ile Pro Ser		
	325	330	335
Asn Phe Ala Lys Arg Val Ala Asp Gln Leu Ile Ser Thr Gly Gln Val			
	340	345	350
Thr Gln Pro Met Ile Gly Val Gln Val Gly Thr Asp Asn Ser Val Thr			
	355	360	365
Gly Ala Val Ile Ala Ser Val Gln Asp Gly Gly Pro Ala Ala Asp Ala			
	370	375	380
Gly Leu Gln Pro Gly Asp Ile Val Thr Lys Leu Asn Asp Arg Val Ile			
	385	390	395
Asp Ser Pro Asp Ser Leu Ile Ala Ala Val Arg Ser His Asp Phe Gly			
	405	410	415
Glu Thr Val Thr Leu Thr Ile Thr Gln Pro Asp Thr Ser Gln Ser Arg			
	420	425	430
Glu Val Glu Val Thr Leu Thr Ser Glu			
	435	440	

<210> 57  
 <211> 518  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (1)..(495)  
 <223> RXA02834

<400> 57	
gat tct aaa ggt aga agt gtt gac ttt aaa aat acc att atc atc atg	48
Asp Ser Lys Gly Arg Ser Val Asp Phe Lys Asn Thr Ile Ile Ile Met	
1 5 10 15	
act agt aat att ggt tca caa gta tta ctt gaa aat gta aaa gat gct	96
Thr Ser Asn Ile Gly Ser Gln Val Leu Glu Asn Val Lys Asp Ala	
20 25 30	
ggt gaa att agt gat gat aca gag aaa gca gtt atg gac agt cta cat	144
Gly Glu Ile Ser Asp Asp Thr Glu Lys Ala Val Met Asp Ser Leu His	
35 40 45	
gca tac ttc aaa cct gaa ata tta aat cgt atg gat gac atc gtg tta	192
Ala Tyr Phe Lys Pro Glu Ile Leu Asn Arg Met Asp Asp Ile Val Leu	
50 55 60	
ttt aaa cca tta tca gtt aat gat atg agt atg att gta gat aaa att	240
Phe Lys Pro Leu Ser Val Asn Asp Met Ser Met Ile Val Asp Lys Ile	
65 70 75 80	
tta aca caa tta aat atg aga tta tta gat caa cat atc tca att gaa	288
Leu Thr Gln Leu Asn Met Arg Leu Leu Asp Gln His Ile Ser Ile Glu	
85 90 95	

gtg aca gaa gaa gcg aaa aaa tgg cta ggt gaa gaa gcg tat gaa cca 336  
 Val Thr Glu Glu Ala Lys Lys Trp Leu Gly Glu Glu Ala Tyr Glu Pro  
                   100                  105                  110

caa ttt ggt gca aga cca tta aaa cgc ttt gtt caa cga caa ata gaa 384  
 Gln Phe Gly Ala Arg Pro Leu Lys Arg Phe Val Gln Arg Gln Ile Glu  
                   115                  120                  125

act cca att gca cgt atg atg att aaa gaa agt cta cct gaa ggt aca 432  
 Thr Pro Ile Ala Arg Met Met Ile Lys Glu Ser Leu Pro Glu Gly Thr  
                   130                  135                  140

ata att aaa gta gat tta aat gac aat aaa gaa ctt gat ttt aag gtt 480  
 Ile Ile Lys Val Asp Leu Asn Asp Asn Lys Glu Leu Asp Phe Lys Val  
                   145                  150                  155                  160

gtt aaa cct acg tct taatctagca aaatattaat ttg 518  
 Val Lys Pro Thr Ser  
                   165

<210> 58  
 <211> 165  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 58  
 Asp Ser Lys Gly Arg Ser Val Asp Phe Lys Asn Thr Ile Ile Ile Met  
   1                  5                  10                  15

Thr Ser Asn Ile Gly Ser Gln Val Leu Leu Glu Asn Val Lys Asp Ala  
                   20                  25                  30

Gly Glu Ile Ser Asp Asp Thr Glu Lys Ala Val Met Asp Ser Leu His  
                   35                  40                  45

Ala Tyr Phe Lys Pro Glu Ile Leu Asn Arg Met Asp Asp Ile Val Leu  
                   50                  55                  60

Phe Lys Pro Leu Ser Val Asn Asp Met Ser Met Ile Val Asp Lys Ile  
   65                  70                  75                  80

Leu Thr Gln Leu Asn Met Arg Leu Leu Asp Gln His Ile Ser Ile Glu  
                   85                  90                  95

Val Thr Glu Glu Ala Lys Lys Trp Leu Gly Glu Glu Ala Tyr Glu Pro  
                   100                  105                  110

Gln Phe Gly Ala Arg Pro Leu Lys Arg Phe Val Gln Arg Gln Ile Glu  
                   115                  120                  125

Thr Pro Ile Ala Arg Met Met Ile Lys Glu Ser Leu Pro Glu Gly Thr  
                   130                  135                  140

Ile Ile Lys Val Asp Leu Asn Asp Asn Lys Glu Leu Asp Phe Lys Val  
   145                  150                  155                  160

Val Lys Pro Thr Ser  
                   165

```
<220>
<221> CDS
<222> (101)..(1291)
<223> RXA00112
```

85

gaa gtg gac gcc ccc gaa atc aac gtc acc aac gtt gac cta gtc gaa 691  
 Glu Val Asp Ala Pro Glu Ile Asn Val Thr Asn Val Asp Leu Val Glu  
 185 190 195

gca atg cgc ccg tcc gtc atc cac gtg atg ggt gac gcc caa gaa tgc 739  
 Ala Met Arg Pro Ser Val Ile His Val Met Gly Asp Ala Gln Glu Cys  
 200 205 210

agc cgc cga ctc atg ggt tct ggc ttt gtg gca tcc ccc gac tac gtt 787  
 Ser Arg Arg Leu Met Gly Ser Gly Phe Val Ala Ser Pro Asp Tyr Val  
 215 220 225

gtg acc aac gcc cac gtt gtt gca ggt acc tcc acc gtc agc ctg gat 835  
 Val Thr Asn Ala His Val Val Ala Gly Thr Ser Thr Val Ser Leu Asp  
 230 235 240 245

acc atg atc gga acc cgc tcc gca gag gta gtg ttc tac gac ccg aac 883  
 Thr Met Ile Gly Thr Arg Ser Ala Glu Val Val Phe Tyr Asp Pro Asn  
 250 255 260

ctg gac atc gca gtc ctt tac agc cct gac ctc ggc ttg gat cca ctg 931  
 Leu Asp Ile Ala Val Leu Tyr Ser Pro Asp Leu Gly Leu Asp Pro Leu  
 265 270 275

ccg tgg gca tcc act ccg cta gac act ggc gat gaa gca atc gtc atg 979  
 Pro Trp Ala Ser Thr Pro Leu Asp Thr Gly Asp Glu Ala Ile Val Met  
 280 285 290

gga ttc cca cag tcc gga cct ttc aac gcc tcc cca gcc agg gtc cgc 1027  
 Gly Phe Pro Gln Ser Gly Pro Phe Asn Ala Ser Pro Ala Arg Val Arg  
 295 300 305

gaa cgc atc atg atc acc ggc agc aac att tac gcc aac ggc cag cac 1075  
 Glu Arg Ile Met Ile Thr Gly Ser Asn Ile Tyr Ala Asn Gly Gln His  
 310 315 320 325

gaa cgc gaa gcc tat tca gtc cgc gga tcc atc caa tct gga aac tcc 1123  
 Glu Arg Glu Ala Tyr Ser Val Arg Gly Ser Ile Gln Ser Gly Asn Ser  
 330 335 340

ggc ggc cca atg acc aac gaa atg ggt gaa gtg gtt ggt gtt gtc ttc 1171  
 Gly Gly Pro Met Thr Asn Glu Met Gly Glu Val Val Gly Val Val Phe  
 345 350 355

ggc gca gcg atc gac ggc tcc gat acc ggt tac gtt ctc act gcc gaa 1219  
 Gly Ala Ala Ile Asp Gly Ser Asp Thr Gly Tyr Val Leu Thr Ala Glu  
 360 365 370

gag gta cag gag cgg atc ggc gac atc acc gcg ctg act cag cct gtc 1267  
 Glu Val Gln Glu Arg Ile Gly Asp Ile Thr Ala Leu Thr Gln Pro Val  
 375 380 385

gat acg atg cag tgc gcg gtt tct tagtcgtcgg gagctaggac cag 1314  
 Asp Thr Met Gln Cys Ala Val Ser  
 390 395

<210> 60  
 <211> 397  
 <212> PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 60

Leu Ser Pro Ser Leu Val Val Asp Ala Val Ile Val Leu Val Met Ala  
 1 5 10 15

Phe Ala Leu Trp Gly Gly Trp Arg Gln Gly Ala Phe Thr Ser Leu Leu  
 20 25 30

Ser Thr Val Gly Val Val Ser Gly Leu Val Val Gly Ala Ala Ala Ala  
 35 40 45

Pro Phe Val Met Gly Leu Thr Asp Ser Thr Ala Leu Arg Phe Leu Leu  
 50 55 60

Ala Ile Gly Thr Val Val Leu Leu Val Gly Leu Gly Asn Leu Ile Gly  
 65 70 75 80

Ala His Leu Gly Ala Ala Ile Arg Asp Asn Ile Lys Phe Arg Ser Ser  
 85 90 95

Arg Thr Leu Asp Ser Gly Leu Gly Ala Ile Phe Gln Val Leu Ala Thr  
 100 105 110

Leu Ile Val Val Trp Leu Val Ala Ile Pro Leu Ala Thr Gly Leu Pro  
 115 120 125

Gly Thr Val Ala Ser Gly Ile Arg Asp Ser Arg Ile Leu Gly Phe Val  
 130 135 140

Asp Lys Tyr Thr Pro Gln Gly Leu Asp Thr Leu Pro Ser Lys Ile Ala  
 145 150 155 160

Ala Met Leu Ser Glu Ser Gly Leu Pro Pro Leu Ile Ser Pro Phe Thr  
 165 170 175

Gly Gly Ser Ser Val Glu Val Asp Ala Pro Glu Ile Asn Val Thr Asn  
 180 185 190

Val Asp Leu Val Glu Ala Met Arg Pro Ser Val Ile His Val Met Gly  
 195 200 205

Asp Ala Gln Glu Cys Ser Arg Arg Leu Met Gly Ser Gly Phe Val Ala  
 210 215 220

Ser Pro Asp Tyr Val Val Thr Asn Ala His Val Val Ala Gly Thr Ser  
 225 230 235 240

Thr Val Ser Leu Asp Thr Met Ile Gly Thr Arg Ser Ala Glu Val Val  
 245 250 255

Phe Tyr Asp Pro Asn Leu Asp Ile Ala Val Leu Tyr Ser Pro Asp Leu  
 260 265 270

Gly Leu Asp Pro Leu Pro Trp Ala Ser Thr Pro Leu Asp Thr Gly Asp  
 275 280 285

Glu Ala Ile Val Met Gly Phe Pro Gln Ser Gly Pro Phe Asn Ala Ser  
 290 295 300

Pro Ala Arg Val Arg Glu Arg Ile Met Ile Thr Gly Ser Asn Ile Tyr

305		310		315		320
Ala Asn Gly Gln His Glu Arg Glu Ala Tyr Ser Val Arg Gly Ser Ile						
	325			330		335
Gln Ser Gly Asn Ser Gly Gly Pro Met Thr Asn Glu Met Gly Glu Val						
	340			345		350
Val Gly Val Val Phe Gly Ala Ala Ile Asp Gly Ser Asp Thr Gly Tyr						
	355			360		365
Val Leu Thr Ala Glu Glu Val Gln Glu Arg Ile Gly Asp Ile Thr Ala						
	370			375		380
Leu Thr Gln Pro Val Asp Thr Met Gln Cys Ala Val Ser						
	385			390		395

<210> 61  
 <211> 729  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(706)  
 <223> RXA00566

<400> 61  
 cttgttgacc acgtgattac cttggctgaa ggcccaatca gcaactaggc gcacggaaaa 60  
 ctttaaagga gagaataaga ttatgagcag cggattccaa atg cca acg tcc cgt 115  
 Met Pro Thr Ser Arg  
 1 5  
 tac gtg ctg cct tcc ttc att gag cag tcc gca tac ggc acc aaa gag 163  
 Tyr Val Leu Pro Ser Phe Ile Glu Gln Ser Ala Tyr Gly Thr Lys Glu  
 10 15 20  
 acc aac cct tac gca aaa ctc ttt gaa gag cgc atc atc ttc ctg ggc 211  
 Thr Asn Pro Tyr Ala Lys Leu Phe Glu Glu Arg Ile Ile Phe Leu Gly  
 25 30 35  
 acc cag gtc gac gac acc tct gca aac gac atc atg gcg cag ctc ctt 259  
 Thr Gln Val Asp Asp Thr Ser Ala Asn Asp Ile Met Ala Gln Leu Leu  
 40 45 50  
 gtc ctc gaa ggc atg gac cca gac cgc gat atc acc ctg tac atc aac 307  
 Val Leu Glu Gly Met Asp Pro Asp Arg Asp Ile Thr Leu Tyr Ile Asn  
 55 60 65  
 tca cct ggt gga tcc ttc acc gcg ttg atg gca att tac gac acc atg 355  
 Ser Pro Gly Gly Ser Phe Thr Ala Leu Met Ala Ile Tyr Asp Thr Met  
 70 75 80 85  
 cag tac gtc cgc cca gac gtt cag acc gtc tgc ctt ggt cag gca gca 403  
 Gln Tyr Val Arg Pro Asp Val Gln Thr Val Cys Leu Gly Gln Ala Ala  
 90 95 100  
 tcc gca gcc gca gtt ctt ctt gca gcc ggt gca cca ggt aag cgc gct 451  
 Ser Ala Ala Ala Val Leu Leu Ala Ala Gly Ala Pro Gly Lys Arg Ala



105	110	115	
ggt ctt cct aac tcc cgc gtg ctg atc cac cag cca gca acc cag ggc			499
Val Leu Pro Asn Ser Arg Val Leu Ile His Gln Pro Ala Thr Gln Gly			
120	125	130	
acc cag ggt cag gtt tct gac ctc gag atc cag gca gct gaa atc gag			547
Thr Gln Gly Gln Val Ser Asp Leu Glu Ile Gln Ala Ala Glu Ile Glu			
135	140	145	
cgc atg cgt cgt ttg atg gaa acc acc ttg gct gag cac acc ggc aag			595
Arg Met Arg Arg Leu Met Glu Thr Thr Leu Ala Glu His Thr Gly Lys			
150	155	160	165
acc gcg gag cag atc cgc atc gat acc gac cgt gac aag atc ctc acc			643
Thr Ala Glu Gln Ile Arg Ile Asp Thr Asp Arg Asp Lys Ile Leu Thr			
170	175	180	
gct gag gaa gca ctc gag tat ggc atc gtt gac cag gtc ttc gat tac			691
Ala Glu Glu Ala Leu Glu Tyr Gly Ile Val Asp Gln Val Phe Asp Tyr			
185	190	195	
cgc aag ctc aag cgc tagagttttt taaagatttcg ggt			729
Arg Lys Leu Lys Arg			
200			

&lt;210&gt; 62

&lt;211&gt; 202

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 62

Met Pro Thr Ser Arg Tyr Val Leu Pro Ser Phe Ile Glu Gln Ser Ala			
1	5	10	15
Tyr Gly Thr Lys Glu Thr Asn Pro Tyr Ala Lys Leu Phe Glu Glu Arg			
20	25	30	
Ile Ile Phe Leu Gly Thr Gln Val Asp Asp Thr Ser Ala Asn Asp Ile			
35	40	45	
Met Ala Gln Leu Leu Val Leu Glu Gly Met Asp Pro Asp Arg Asp Ile			
50	55	60	
Thr Leu Tyr Ile Asn Ser Pro Gly Gly Ser Phe Thr Ala Leu Met Ala			
65	70	75	80
Ile Tyr Asp Thr Met Gln Tyr Val Arg Pro Asp Val Gln Thr Val Cys			
85	90	95	
Leu Gly Gln Ala Ala Ser Ala Ala Ala Val Leu Leu Ala Ala Gly Ala			
100	105	110	
Pro Gly Lys Arg Ala Val Leu Pro Asn Ser Arg Val Leu Ile His Gln			
115	120	125	
Pro Ala Thr Gln Gly Thr Gln Gly Gln Val Ser Asp Leu Glu Ile Gln			
130	135	140	
Ala Ala Glu Ile Glu Arg Met Arg Arg Leu Met Glu Thr Thr Leu Ala			

<400> 63																
cgaacagagg	cggtttc	catg	gaaataccgc			cgggtagtct			ggtgacattg			aaccaa			atga	60
acgtacccaa	gatttaagaa			tgtaggagtt			gactgttttc			atg	agc	gat	att	cgt	115	
										Met	Ser	Asp	Ile	Arg		
										1				5		
atg	gca	gcc	cag	ggt	ggg	cct	ggt	ttc	gga	aat	gac	gtc	ttt	gat	cgc	163
Met	Ala	Ala	Gln	Gly	Gly	Pro	Gly	Phe	Gly	Asn	Asp	Val	Phe	Asp	Arg	
				10					15					20		
ctc	atg	agt	gag	cgc	atc	att	ttc	ctg	gga	agc	cag	gta	gac	gat	gag	211
Leu	Met	Ser	Glu	Arg	Ile	Ile	Phe	Leu	Gly	Ser	Gln	Val	Asp	Asp	Glu	
			25					30					35			
att	gca	aac	aag	cta	tgc	gct	cag	atc	ctg	ctg	ctg	tcc	gct	gag	gat	259
Ile	Ala	Asn	Lys	Leu	Cys	Ala	Gln	Ile	Leu	Leu	Leu	Ser	Ala	Glu	Asp	
		40					45					50				
cca	acc	agg	gac	atc	tcc	ctg	tac	att	aac	tcc	cca	ggg	ggc	tcc	gtc	307
Pro	Thr	Arg	Asp	Ile	Ser	Leu	Tyr	Ile	Asn	Ser	Pro	Gly	Gly	Ser	Val	
	55					60					65					
acc	gca	ggc	atg	gca	att	tat	gac	acc	atg	aaa	tac	tcc	cca	tgc	gac	355
Thr	Ala	Gly	Met	Ala	Ile	Tyr	Asp	Thr	Met	Lys	Tyr	Ser	Pro	Cys	Asp	
70					75					80					85	
atc	gca	acc	tac	ggc	atg	ggc	ctg	gca	gca	tcc	atg	ggc	cag	ttc	ctg	403
Ile	Ala	Thr	Tyr	Gly	Met	Gly	Leu	Ala	Ala	Ser	Met	Gly	Gln	Phe	Leu	
				90					95					100		
ctt	tcc	ggt	ggc	act	aag	ggc	aag	cgt	ttc	gca	ttg	cca	cac	gca	cgt	451
Leu	Ser	Gly	Gly	Thr	Lys	Gly	Lys	Arg	Phe	Ala	Leu	Pro	His	Ala	Arg	
			105					110					115			
atc	atg	atg	cac	cag	cct	tcc	gct	ggt	gtt	ggt	ggt	acc	gca	gct	gat	499
Ile	Met	Met	His	Gln	Pro	Ser	Ala	Gly	Val	Gly	Gly	Thr	Ala	Ala	Asp	
		120					125					130				

atc gct atc cag gct gag cag ttc gca gcc acc aag cgt gaa atg gcc 547  
 Ile Ala Ile Gln Ala Glu Gln Phe Ala Ala Thr Lys Arg Glu Met Ala  
 135 140 145

cag ctg atc gct gag cac acc ggc cag acc ttt gag cag atc tcc aag 595  
 Gln Leu Ile Ala Glu His Thr Gly Gln Thr Phe Glu Gln Ile Ser Lys  
 150 155 160 165

gat tcc gat cgt gac cgc tgg ttc act gca cag gaa gct aag gat tac 643  
 Asp Ser Asp Arg Asp Arg Trp Phe Thr Ala Gln Glu Ala Lys Asp Tyr  
 170 175 180

gga ctt gtt gac cac gtg att acc ttg gct gaa ggc cca atc agc aac 691  
 Gly Leu Val Asp His Val Ile Thr Leu Ala Glu Gly Pro Ile Ser Asn  
 185 190 195

taggcgcacg gaaaacttta aag 714

<210> 64  
 <211> 197  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 64  
 Met Ser Asp Ile Arg Met Ala Ala Gln Gly Gly Pro Gly Phe Gly Asn  
 1 5 10 15

Asp Val Phe Asp Arg Leu Met Ser Glu Arg Ile Ile Phe Leu Gly Ser  
 20 25 30

Gln Val Asp Asp Glu Ile Ala Asn Lys Leu Cys Ala Gln Ile Leu Leu  
 35 40 45

Leu Ser Ala Glu Asp Pro Thr Arg Asp Ile Ser Leu Tyr Ile Asn Ser  
 50 55 60

Pro Gly Gly Ser Val Thr Ala Gly Met Ala Ile Tyr Asp Thr Met Lys  
 65 70 75 80

Tyr Ser Pro Cys Asp Ile Ala Thr Tyr Gly Met Gly Leu Ala Ala Ser  
 85 90 95

Met Gly Gln Phe Leu Leu Ser Gly Gly Thr Lys Gly Lys Arg Phe Ala  
 100 105 110

Leu Pro His Ala Arg Ile Met Met His Gln Pro Ser Ala Gly Val Gly  
 115 120 125

Gly Thr Ala Ala Asp Ile Ala Ile Gln Ala Glu Gln Phe Ala Ala Thr  
 130 135 140

Lys Arg Glu Met Ala Gln Leu Ile Ala Glu His Thr Gly Gln Thr Phe  
 145 150 155 160

Glu Gln Ile Ser Lys Asp Ser Asp Arg Asp Arg Trp Phe Thr Ala Gln  
 165 170 175

Glu Ala Lys Asp Tyr Gly Leu Val Asp His Val Ile Thr Leu Ala Glu  
 180 185 190

Gly Pro Ile Ser Asn  
195

<210> 65  
<211> 1875  
<212> DNA  
<213> Corynebacterium glutamicum

<220>  
<221> CDS  
<222> (101)..(1852)  
<223> RXN03094

<400> 65  
atagatatta gagagttaaa taatggcgct tgacctgcag gaaattgaga tcaacactga 60  
ttgtgtaggt tggcgcccaa caaagaaagg gcgttgaaag atg agt tca ttc aat 115  
Met Ser Ser Phe Asn  
1 5  
cca act acc aaa acc aat gaa gcc atg cag gct gct ctt cag cag gca 163  
Pro Thr Thr Lys Thr Asn Glu Ala Met Gln Ala Ala Leu Gln Gln Ala  
10 15 20  
tcc tcg gct ggc aac cct gat att cgt cca gct cac ctg ttg gct gcc 211  
Ser Ser Ala Gly Asn Pro Asp Ile Arg Pro Ala His Leu Leu Ala Ala  
25 30 35  
atc ttg gag caa act gat ggc gta gca gcg cca gtc ctc atg gct act 259  
Ile Leu Glu Gln Thr Asp Gly Val Ala Ala Pro Val Leu Met Ala Thr  
40 45 50  
ggg gtg gat cct aag gag atc ctc gca gag gcc aag aag ttg gtt gct 307  
Gly Val Asp Pro Lys Glu Ile Leu Ala Glu Ala Lys Lys Leu Val Ala  
55 60 65  
tct tac ccc aag gct tct ggc gcc aat atg gct aat cca aac ttc aac 355  
Ser Tyr Pro Lys Ala Ser Gly Ala Asn Met Ala Asn Pro Asn Phe Asn  
70 75 80 85  
cgg gat gcc ctc aat gcg ttc act gca gct cag gag ctt gcc ggt gag 403  
Arg Asp Ala Leu Asn Ala Phe Thr Ala Ala Gln Glu Leu Ala Gly Glu  
90 95 100  
ttg ggc gat gag tac gtc tca acc gaa gta ctt ctt gcc ggt atc gct 451  
Leu Gly Asp Glu Tyr Val Ser Thr Glu Val Leu Leu Ala Gly Ile Ala  
105 110 115  
cgc gga aag tct gat gct gcg gat ctg ttg acc aac aag ggt gca acc 499  
Arg Gly Lys Ser Asp Ala Ala Asp Leu Leu Thr Asn Lys Gly Ala Thr  
120 125 130  
tat gac gcc atc aaa gag gct ttc cct tcg gtt cgt gga tct cag cgt 547  
Tyr Asp Ala Ile Lys Glu Ala Phe Pro Ser Val Arg Gly Ser Gln Arg  
135 140 145  
gtc acc act cag gat cca gag gga cag ttc cag gct ttg gaa aag tac 595  
Val Thr Thr Gln Asp Pro Glu Gly Gln Phe Gln Ala Leu Glu Lys Tyr  
150 155 160 165

tcc act gac ctg acc aag ctt gct cgt gaa ggc aag att gat cct gtt	643
Ser Thr Asp Leu Thr Lys Leu Ala Arg Glu Gly Lys Ile Asp Pro Val	
170 175 180	
att ggc cgt gac cag gaa att cgt cgc gtc gtt cag gtg ctt agc cgt	691
Ile Gly Arg Asp Gln Glu Ile Arg Arg Val Val Gln Val Leu Ser Arg	
185 190 195	
cgt acc aag aac aac cct gtt ctg atc ggt gag cca ggt gtc ggt aaa	739
Arg Thr Lys Asn Asn Pro Val Leu Ile Gly Glu Pro Gly Val Gly Lys	
200 205 210	
acc gcc atc gtg gaa ggc ctt gca cgc cgc atc gtt gct ggt gac gtt	787
Thr Ala Ile Val Glu Gly Leu Ala Arg Arg Ile Val Ala Gly Asp Val	
215 220 225	
cca gaa tcc ctc aag ggc aaa act ctg atc agt ctt gat ctt ggt tcc	835
Pro Glu Ser Leu Lys Gly Lys Thr Leu Ile Ser Leu Asp Leu Gly Ser	
230 235 240 245	
atg gtt gcc ggc gct aag tat cgc ggt gaa ttc gag gag cga ctg aag	883
Met Val Ala Gly Ala Lys Tyr Arg Gly Glu Phe Glu Glu Arg Leu Lys	
250 255 260	
gct gtt ctg gat gag atc aag gga gct aac ggc gaa gtc gtt acc ttc	931
Ala Val Leu Asp Glu Ile Lys Gly Ala Asn Gly Glu Val Val Thr Phe	
265 270 275	
atc gat gag ctg cac acc atc gtc ggc gct ggt gct tcg ggt gaa tcc	979
Ile Asp Glu Leu His Thr Ile Val Gly Ala Gly Ala Ser Gly Glu Ser	
280 285 290	
gcc atg gat gcc gga aac atg att aag cca ctg ctt gcc cgc ggt gag	1027
Ala Met Asp Ala Gly Asn Met Ile Lys Pro Leu Leu Ala Arg Gly Glu	
295 300 305	
ctg cgc ttg gtt ggt gcc acc acg ctg aat gag tac cgc aag tac atc	1075
Leu Arg Leu Val Gly Ala Thr Thr Leu Asn Glu Tyr Arg Lys Tyr Ile	
310 315 320 325	
gaa aag gac gct gcc ctg gag cgt agg ttc cag cag gtt tat gtc ggt	1123
Glu Lys Asp Ala Ala Leu Glu Arg Arg Phe Gln Gln Val Tyr Val Gly	
330 335 340	
gag cca acg gta gaa gat gcc atc ggt att ctt cgt gga ttg aag gaa	1171
Glu Pro Thr Val Glu Asp Ala Ile Gly Ile Leu Arg Gly Leu Lys Glu	
345 350 355	
cgc tac gag gtc cat cac ggt gtc cgc atc cag gac tcc gca ctg gtc	1219
Arg Tyr Glu Val His His Gly Val Arg Ile Gln Asp Ser Ala Leu Val	
360 365 370	
gcc gca gct gaa ctc tca aac cgc tat atc acc agc cgt ttc ctt cct	1267
Ala Ala Ala Glu Leu Ser Asn Arg Tyr Ile Thr Ser Arg Phe Leu Pro	
375 380 385	
gat aag gct att gac tta gtt gat gag gca gca tca cgc ctg cgc atg	1315
Asp Lys Ala Ile Asp Leu Val Asp Glu Ala Ala Ser Arg Leu Arg Met	
390 395 400 405	
gag att gat tct tca cct cag gaa atc gat gag ctg gag cgt atc gtc	1363

Glu Ile Asp Ser Ser Pro Gln Glu Ile Asp Glu Leu Glu Arg Ile Val  
 410 415 420

cgc cgc ctc gag atc gaa gag atg gcg ctg tcc aag gaa tcc gat gca 1411  
 Arg Arg Leu Glu Ile Glu Glu Met Ala Leu Ser Lys Glu Ser Asp Ala  
 425 430 435

gct tcc aag gaa cgt cta gaa aag ctg cgc tcg gaa ctt gct gat gaa 1459  
 Ala Ser Lys Glu Arg Leu Glu Lys Leu Arg Ser Glu Leu Ala Asp Glu  
 440 445 450

cgc gaa aag ctc tct gag ttg aag gct cgt tgg cag aat gag aaa act 1507  
 Arg Glu Lys Leu Ser Glu Leu Lys Ala Arg Trp Gln Asn Glu Lys Thr  
 455 460 465

gct att gac gat gtc cgg gag atg aaa gaa gag ctg gaa gcg ctg cgt 1555  
 Ala Ile Asp Asp Val Arg Glu Met Lys Glu Glu Leu Glu Ala Leu Arg  
 470 475 480 485

tct gag tcg gat att gca aaa cgt gac ggc aat tat tgt cgt gtc gca 1603  
 Ser Glu Ser Asp Ile Ala Lys Arg Asp Gly Asn Tyr Cys Arg Val Ala  
 490 495 500

aag ctt cgc tac ggc cga atc cct gag ctg gaa aag cag atc gag gat 1651  
 Lys Leu Arg Tyr Gly Arg Ile Pro Glu Leu Glu Lys Gln Ile Glu Asp  
 505 510 515

gca gaa tcc aag gtc gag gtc aat gaa aat gcc atg ctc act gag gag 1699  
 Ala Glu Ser Lys Val Glu Val Asn Glu Asn Ala Met Leu Thr Glu Glu  
 520 525 530

gtc acg cca gac acg atc gcc gat gtg gtt tcc gca tgg acg ggc att 1747  
 Val Thr Pro Asp Thr Ile Ala Asp Val Val Ser Ala Trp Thr Gly Ile  
 535 540 545

cct gca ggc aag atg atg cag ggt gag acc gag aag ctg ctc aac atg 1795  
 Pro Ala Gly Lys Met Met Gln Gly Glu Thr Glu Lys Leu Leu Asn Met  
 550 555 560 565

gag cgc gtc ttg ggc aac cgt gtg gtc ggt cag cta gaa agc ggt aac 1843  
 Glu Arg Val Leu Gly Asn Arg Val Val Gly Gln Leu Glu Ser Gly Asn  
 570 575 580

tgc agt gtc tgacgcggtg cgtcgttcac gcg 1875  
 Cys Ser Val

&lt;210&gt; 66

&lt;211&gt; 584

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 66

Met Ser Ser Phe Asn Pro Thr Thr Lys Thr Asn Glu Ala Met Gln Ala  
 1 5 10 15

Ala Leu Gln Gln Ala Ser Ser Ala Gly Asn Pro Asp Ile Arg Pro Ala  
 20 25 30

His Leu Leu Ala Ala Ile Leu Glu Gln Thr Asp Gly Val Ala Ala Pro

35					40					45					
Val	Leu	Met	Ala	Thr	Gly	Val	Asp	Pro	Lys	Glu	Ile	Leu	Ala	Glu	Ala
50					55					60					
Lys	Lys	Leu	Val	Ala	Ser	Tyr	Pro	Lys	Ala	Ser	Gly	Ala	Asn	Met	Ala
65					70					75					80
Asn	Pro	Asn	Phe	Asn	Arg	Asp	Ala	Leu	Asn	Ala	Phe	Thr	Ala	Ala	Gln
				85					90					95	
Glu	Leu	Ala	Gly	Glu	Leu	Gly	Asp	Glu	Tyr	Val	Ser	Thr	Glu	Val	Leu
			100					105					110		
Leu	Ala	Gly	Ile	Ala	Arg	Gly	Lys	Ser	Asp	Ala	Ala	Asp	Leu	Leu	Thr
		115					120					125			
Asn	Lys	Gly	Ala	Thr	Tyr	Asp	Ala	Ile	Lys	Glu	Ala	Phe	Pro	Ser	Val
	130					135					140				
Arg	Gly	Ser	Gln	Arg	Val	Thr	Thr	Gln	Asp	Pro	Glu	Gly	Gln	Phe	Gln
145					150					155					160
Ala	Leu	Glu	Lys	Tyr	Ser	Thr	Asp	Leu	Thr	Lys	Leu	Ala	Arg	Glu	Gly
			165						170					175	
Lys	Ile	Asp	Pro	Val	Ile	Gly	Arg	Asp	Gln	Glu	Ile	Arg	Arg	Val	Val
		180						185					190		
Gln	Val	Leu	Ser	Arg	Arg	Thr	Lys	Asn	Asn	Pro	Val	Leu	Ile	Gly	Glu
	195						200					205			
Pro	Gly	Val	Gly	Lys	Thr	Ala	Ile	Val	Glu	Gly	Leu	Ala	Arg	Arg	Ile
	210					215					220				
Val	Ala	Gly	Asp	Val	Pro	Glu	Ser	Leu	Lys	Gly	Lys	Thr	Leu	Ile	Ser
225					230					235					240
Leu	Asp	Leu	Gly	Ser	Met	Val	Ala	Gly	Ala	Lys	Tyr	Arg	Gly	Glu	Phe
			245					250					255		
Glu	Glu	Arg	Leu	Lys	Ala	Val	Leu	Asp	Glu	Ile	Lys	Gly	Ala	Asn	Gly
			260					265					270		
Glu	Val	Val	Thr	Phe	Ile	Asp	Glu	Leu	His	Thr	Ile	Val	Gly	Ala	Gly
	275						280					285			
Ala	Ser	Gly	Glu	Ser	Ala	Met	Asp	Ala	Gly	Asn	Met	Ile	Lys	Pro	Leu
	290					295					300				
Leu	Ala	Arg	Gly	Glu	Leu	Arg	Leu	Val	Gly	Ala	Thr	Thr	Leu	Asn	Glu
305					310					315					320
Tyr	Arg	Lys	Tyr	Ile	Glu	Lys	Asp	Ala	Ala	Leu	Glu	Arg	Arg	Phe	Gln
			325					330						335	
Gln	Val	Tyr	Val	Gly	Glu	Pro	Thr	Val	Glu	Asp	Ala	Ile	Gly	Ile	Leu
		340						345					350		
Arg	Gly	Leu	Lys	Glu	Arg	Tyr	Glu	Val	His	His	Gly	Val	Arg	Ile	Gln
	355						360					365			

Asp Ser Ala Leu Val Ala Ala Ala Glu Leu Ser Asn Arg Tyr Ile Thr  
 370 375 380  
 Ser Arg Phe Leu Pro Asp Lys Ala Ile Asp Leu Val Asp Glu Ala Ala  
 385 390 395 400  
 Ser Arg Leu Arg Met Glu Ile Asp Ser Ser Pro Gln Glu Ile Asp Glu  
 405 410 415  
 Leu Glu Arg Ile Val Arg Arg Leu Glu Ile Glu Glu Met Ala Leu Ser  
 420 425 430  
 Lys Glu Ser Asp Ala Ala Ser Lys Glu Arg Leu Glu Lys Leu Arg Ser  
 435 440 445  
 Glu Leu Ala Asp Glu Arg Glu Lys Leu Ser Glu Leu Lys Ala Arg Trp  
 450 455 460  
 Gln Asn Glu Lys Thr Ala Ile Asp Asp Val Arg Glu Met Lys Glu Glu  
 465 470 475 480  
 Leu Glu Ala Leu Arg Ser Glu Ser Asp Ile Ala Lys Arg Asp Gly Asn  
 485 490 495  
 Tyr Cys Arg Val Ala Lys Leu Arg Tyr Gly Arg Ile Pro Glu Leu Glu  
 500 505 510  
 Lys Gln Ile Glu Asp Ala Glu Ser Lys Val Glu Val Asn Glu Asn Ala  
 515 520 525  
 Met Leu Thr Glu Glu Val Thr Pro Asp Thr Ile Ala Asp Val Val Ser  
 530 535 540  
 Ala Trp Thr Gly Ile Pro Ala Gly Lys Met Met Gln Gly Glu Thr Glu  
 545 550 555 560  
 Lys Leu Leu Asn Met Glu Arg Val Leu Gly Asn Arg Val Val Gly Gln  
 565 570 575  
 Leu Glu Ser Gly Asn Cys Ser Val  
 580

<210> 67  
 <211> 1816  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1816)  
 <223> FRXA01668

<400> 67  
 atagatatta gagagttaaa taatggcgct tgacctgcag gaaattgaga tcaacactga 60  
 ttgtgtaggt tggcgcccaa caaagaaagg gcgttgaaag atg agt tca ttc aat 115  
 Met Ser Ser Phe Asn  
 1 5



cca act acc aaa acc aat gaa gcc atg cag gct gct ctt cag cag gca	163
Pro Thr Thr Lys Thr Asn Glu Ala Met Gln Ala Ala Leu Gln Gln Ala	
10 15 20	
tcc tcg gct ggc aac cct gat att cgt cca gct cac ctg ttg gct gcc	211
Ser Ser Ala Gly Asn Pro Asp Ile Arg Pro Ala His Leu Leu Ala Ala	
25 30 35	
atc ttg gag caa act gat ggc gta gca gcg cca gtc ctc atg gct act	259
Ile Leu Glu Gln Thr Asp Gly Val Ala Ala Pro Val Leu Met Ala Thr	
40 45 50	
ggg gtg gat cct aag gag atc ctc gca gag gcc aag aag ttg gtt gct	307
Gly Val Asp Pro Lys Glu Ile Leu Ala Glu Ala Lys Lys Leu Val Ala	
55 60 65	
tct tac ccc aag gct tct ggc gcc aat atg gct aat cca aac ttc aac	355
Ser Tyr Pro Lys Ala Ser Gly Ala Asn Met Ala Asn Pro Asn Phe Asn	
70 75 80 85	
cgg gat gcc ctc aat gcg ttc act gca gct cag gag ctt gcc ggt gag	403
Arg Asp Ala Leu Asn Ala Phe Thr Ala Ala Gln Glu Leu Ala Gly Glu	
90 95 100	
ttg ggc gat gag tac gtc tca acc gaa gta ctt ctt gcc ggt atc gct	451
Leu Gly Asp Glu Tyr Val Ser Thr Glu Val Leu Leu Ala Gly Ile Ala	
105 110 115	
cgc gga aag tct gat gct gcg gat ctg ttg acc aac aag ggt gca acc	499
Arg Gly Lys Ser Asp Ala Ala Asp Leu Leu Thr Asn Lys Gly Ala Thr	
120 125 130	
tat gac gcc atc aaa gag gct ttc cct tcg gtt cgt gga tct cag cgt	547
Tyr Asp Ala Ile Lys Glu Ala Phe Pro Ser Val Arg Gly Ser Gln Arg	
135 140 145	
gtc acc act cag gat cca gag gga cag ttc cag gct ttg gaa aag tac	595
Val Thr Thr Gln Asp Pro Glu Gly Gln Phe Gln Ala Leu Glu Lys Tyr	
150 155 160 165	
tcc act gac ctg acc aag ctt gct cgt gaa ggc aag att gat cct gtt	643
Ser Thr Asp Leu Thr Lys Leu Ala Arg Glu Gly Lys Ile Asp Pro Val	
170 175 180	
att ggc cgt gac cag gaa att cgt cgc gtc gtt cag gtg ctt agc cgt	691
Ile Gly Arg Asp Gln Glu Ile Arg Arg Val Val Gln Val Leu Ser Arg	
185 190 195	
cgt acc aag aac aac cct gtt ctg atc ggt gag cca ggt gtc ggt aaa	739
Arg Thr Lys Asn Asn Pro Val Leu Ile Gly Glu Pro Gly Val Gly Lys	
200 205 210	
acc gcc atc gtg gaa ggc ctt gca cgc cgc atc gtt gct ggt gac gtt	787
Thr Ala Ile Val Glu Gly Leu Ala Arg Arg Ile Val Ala Gly Asp Val	
215 220 225	
cca gaa tcc ctc aag ggc aaa act ctg atc agt ctt gat ctt ggt tcc	835
Pro Glu Ser Leu Lys Gly Lys Thr Leu Ile Ser Leu Asp Leu Gly Ser	
230 235 240 245	
atg gtt gcc ggc gct aag tat cgc ggt gaa ttc gag gag cga ctg aag	883

Met	Val	Ala	Gly	Ala	Lys	Tyr	Arg	Gly	Glu	Phe	Glu	Glu	Arg	Leu	Lys		
				250					255					260			
gct	gtt	ctg	gat	gag	atc	aag	gga	gct	aac	ggc	gaa	gtc	gtt	acc	ttc	931	
Ala	Val	Leu	Asp	Glu	Ile	Lys	Gly	Ala	Asn	Gly	Glu	Val	Val	Thr	Phe		
			265				270						275				
atc	gat	gag	ctg	cac	acc	atc	gtc	ggc	gct	ggg	gct	tcg	ggg	gaa	tcc	979	
Ile	Asp	Glu	Leu	His	Thr	Ile	Val	Gly	Ala	Gly	Ala	Ser	Gly	Glu	Ser		
		280					285					290					
gcc	atg	gat	gcc	gga	aac	atg	att	aag	cca	ctg	ctt	gcc	cgc	ggg	gag	1027	
Ala	Met	Asp	Ala	Gly	Asn	Met	Ile	Lys	Pro	Leu	Leu	Ala	Arg	Gly	Glu		
	295					300					305						
ctg	cgc	ttg	gtt	ggg	gcc	acc	acg	ctg	aat	gag	tac	cgc	aag	tac	atc	1075	
Leu	Arg	Leu	Val	Gly	Ala	Thr	Thr	Leu	Asn	Glu	Tyr	Arg	Lys	Tyr	Ile		
310					315					320					325		
gaa	aag	gac	gct	gcc	ctg	gag	cgt	agg	ttc	cag	cag	gtt	tat	gtc	ggg	1123	
Glu	Lys	Asp	Ala	Ala	Leu	Glu	Arg	Arg	Phe	Gln	Gln	Val	Tyr	Val	Gly		
				330				335						340			
gag	cca	acg	gta	gaa	gat	gcc	atc	ggg	att	ctt	cgt	gga	ttg	aag	gaa	1171	
Glu	Pro	Thr	Val	Glu	Asp	Ala	Ile	Gly	Ile	Leu	Arg	Gly	Leu	Lys	Glu		
			345					350					355				
cgc	tac	gag	gtc	cat	cac	ggg	gtc	cgc	atc	cag	gac	tcc	gca	ctg	gtc	1219	
Arg	Tyr	Glu	Val	His	His	Gly	Val	Arg	Ile	Gln	Asp	Ser	Ala	Leu	Val		
		360					365					370					
gcc	gca	gct	gaa	ctc	tca	aac	cgc	tat	atc	acc	agc	cgt	ttc	ctt	cct	1267	
Ala	Ala	Ala	Glu	Leu	Ser	Asn	Arg	Tyr	Ile	Thr	Ser	Arg	Phe	Leu	Pro		
	375					380					385						
gat	aag	gct	att	gac	tta	gtt	gat	gag	gca	gca	tca	cgc	ctg	cgc	atg	1315	
Asp	Lys	Ala	Ile	Asp	Leu	Val	Asp	Glu	Ala	Ala	Ser	Arg	Leu	Arg	Met		
390					395				400						405		
gag	att	gat	tct	tca	cct	cag	gaa	atc	gat	gag	ctg	gag	cgt	atc	gtc	1363	
Glu	Ile	Asp	Ser	Ser	Pro	Gln	Glu	Ile	Asp	Glu	Leu	Glu	Arg	Ile	Val		
				410				415						420			
cgc	cgc	ctc	gag	atc	gaa	gag	atg	gcg	ctg	tcc	aag	gaa	tcc	gat	gca	1411	
Arg	Arg	Leu	Glu	Ile	Glu	Glu	Met	Ala	Leu	Ser	Lys	Glu	Ser	Asp	Ala		
			425				430						435				
gct	tcc	aag	gaa	cgt	cta	gaa	aag	ctg	cgc	tcg	gaa	ctt	gct	gat	gaa	1459	
Ala	Ser	Lys	Glu	Arg	Leu	Glu	Lys	Leu	Arg	Ser	Glu	Leu	Ala	Asp	Glu		
		440					445					450					
cgc	gaa	aag	ctc	tct	gag	ttg	aag	gct	cgt	tgg	cag	aat	gag	aaa	act	1507	
Arg	Glu	Lys	Leu	Ser	Glu	Leu	Lys	Ala	Arg	Trp	Gln	Asn	Glu	Lys	Thr		
	455					460					465						
gct	att	gac	gat	gtc	cgg	gag	atg	aaa	gaa	gag	ctg	gaa	gcg	ctg	cgt	1555	
Ala	Ile	Asp	Asp	Val	Arg	Glu	Met	Lys	Glu	Glu	Leu	Glu	Ala	Leu	Arg		
470					475					480					485		
tct	gag	tcg	gat	att	gca	aaa	cgt	gac	ggc	aat	tat	tgt	cgt	gtc	gca	1603	
Ser	Glu	Ser	Asp	Ile	Ala	Lys	Arg	Asp	Gly	Asn	Tyr	Cys	Arg	Val	Ala		

490	495	500	
aag ctt cgc tac ggc cga atc cct gag ctg gaa aag cag atc gag gat			1651
Lys Leu Arg Tyr Gly Arg Ile Pro Glu Leu Glu Lys Gln Ile Glu Asp			
505	510	515	
gca gaa tcc aag gtc gag gtc aat gaa aat gcc atg ctc act gag gag			1699
Ala Glu Ser Lys Val Glu Val Asn Glu Asn Ala Met Leu Thr Glu Glu			
520	525	530	
gtc acg cca gac acg atc gcc gat gtg gtt tcc gca tgg acg ggc att			1747
Val Thr Pro Asp Thr Ile Ala Asp Val Val Ser Ala Trp Thr Gly Ile			
535	540	545	
cct gca ggc aag atg atg cag ggt gag acc gag aag ctg ctc aac atg			1795
Pro Ala Gly Lys Met Met Gln Gly Glu Thr Glu Lys Leu Leu Asn Met			
550	555	560	565
gag cgc gtc ttg ggc aac ccg			1816
Glu Arg Val Leu Gly Asn Pro			
570			

&lt;210&gt; 68

&lt;211&gt; 572

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 68

Met Ser Ser Phe Asn Pro Thr Thr Lys Thr Asn Glu Ala Met Gln Ala
1 5 10 15

Ala Leu Gln Gln Ala Ser Ser Ala Gly Asn Pro Asp Ile Arg Pro Ala
20 25 30

His Leu Leu Ala Ala Ile Leu Glu Gln Thr Asp Gly Val Ala Ala Pro
35 40 45

Val Leu Met Ala Thr Gly Val Asp Pro Lys Glu Ile Leu Ala Glu Ala
50 55 60

Lys Lys Leu Val Ala Ser Tyr Pro Lys Ala Ser Gly Ala Asn Met Ala
65 70 75 80

Asn Pro Asn Phe Asn Arg Asp Ala Leu Asn Ala Phe Thr Ala Ala Gln
85 90 95

Glu Leu Ala Gly Glu Leu Gly Asp Glu Tyr Val Ser Thr Glu Val Leu
100 105 110

Leu Ala Gly Ile Ala Arg Gly Lys Ser Asp Ala Ala Asp Leu Leu Thr
115 120 125

Asn Lys Gly Ala Thr Tyr Asp Ala Ile Lys Glu Ala Phe Pro Ser Val
130 135 140

Arg Gly Ser Gln Arg Val Thr Thr Gln Asp Pro Glu Gly Gln Phe Gln
145 150 155 160

Ala Leu Glu Lys Tyr Ser Thr Asp Leu Thr Lys Leu Ala Arg Glu Gly
165 170 175

Lys Ile Asp Pro Val Ile Gly Arg Asp Gln Glu Ile Arg Arg Val Val  
 180 185 190  
 Gln Val Leu Ser Arg Arg Thr Lys Asn Asn Pro Val Leu Ile Gly Glu  
 195 200 205  
 Pro Gly Val Gly Lys Thr Ala Ile Val Glu Gly Leu Ala Arg Arg Ile  
 210 215 220  
 Val Ala Gly Asp Val Pro Glu Ser Leu Lys Gly Lys Thr Leu Ile Ser  
 225 230 235 240  
 Leu Asp Leu Gly Ser Met Val Ala Gly Ala Lys Tyr Arg Gly Glu Phe  
 245 250 255  
 Glu Glu Arg Leu Lys Ala Val Leu Asp Glu Ile Lys Gly Ala Asn Gly  
 260 265 270  
 Glu Val Val Thr Phe Ile Asp Glu Leu His Thr Ile Val Gly Ala Gly  
 275 280 285  
 Ala Ser Gly Glu Ser Ala Met Asp Ala Gly Asn Met Ile Lys Pro Leu  
 290 295 300  
 Leu Ala Arg Gly Glu Leu Arg Leu Val Gly Ala Thr Thr Leu Asn Glu  
 305 310 315 320  
 Tyr Arg Lys Tyr Ile Glu Lys Asp Ala Ala Leu Glu Arg Arg Phe Gln  
 325 330 335  
 Gln Val Tyr Val Gly Glu Pro Thr Val Glu Asp Ala Ile Gly Ile Leu  
 340 345 350  
 Arg Gly Leu Lys Glu Arg Tyr Glu Val His His Gly Val Arg Ile Gln  
 355 360 365  
 Asp Ser Ala Leu Val Ala Ala Ala Glu Leu Ser Asn Arg Tyr Ile Thr  
 370 375 380  
 Ser Arg Phe Leu Pro Asp Lys Ala Ile Asp Leu Val Asp Glu Ala Ala  
 385 390 395 400  
 Ser Arg Leu Arg Met Glu Ile Asp Ser Ser Pro Gln Glu Ile Asp Glu  
 405 410 415  
 Leu Glu Arg Ile Val Arg Arg Leu Glu Ile Glu Glu Met Ala Leu Ser  
 420 425 430  
 Lys Glu Ser Asp Ala Ala Ser Lys Glu Arg Leu Glu Lys Leu Arg Ser  
 435 440 445  
 Glu Leu Ala Asp Glu Arg Glu Lys Leu Ser Glu Leu Lys Ala Arg Trp  
 450 455 460  
 Gln Asn Glu Lys Thr Ala Ile Asp Asp Val Arg Glu Met Lys Glu Glu  
 465 470 475 480  
 Leu Glu Ala Leu Arg Ser Glu Ser Asp Ile Ala Lys Arg Asp Gly Asn  
 485 490 495

Tyr Cys Arg Val Ala Lys Leu Arg Tyr Gly Arg Ile Pro Glu Leu Glu  
 500 505 510  
 Lys Gln Ile Glu Asp Ala Glu Ser Lys Val Glu Val Asn Glu Asn Ala  
 515 520 525  
 Met Leu Thr Glu Glu Val Thr Pro Asp Thr Ile Ala Asp Val Val Ser  
 530 535 540  
 Ala Trp Thr Gly Ile Pro Ala Gly Lys Met Met Gln Gly Glu Thr Glu  
 545 550 555 560  
 Lys Leu Leu Asn Met Glu Arg Val Leu Gly Asn Pro  
 565 570

<210> 69  
 <211> 1401  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1378)  
 <223> RXN01120

<400> 69  
 acaggtaaag cgctaagatg gaacaaccca ttgccaatat tgttggttag agttgtacgc 60  
 agtaaatctt ttcaatcgtg gaagcgggtc tcacagtcta atg gca cgt atg cag 115  
 Met Ala Arg Met Gln  
 1 5  
 gaa agc gcc gat ctg ctc aaa tgt tcc ttc tgc gga aag agc caa aag 163  
 Glu Ser Ala Asp Leu Leu Lys Cys Ser Phe Cys Gly Lys Ser Gln Lys  
 10 15 20  
 cag gta aaa aaa ctc atc gcg ggt ggc gcc gta tat atc tgt gat gag 211  
 Gln Val Lys Lys Leu Ile Ala Gly Gly Ala Val Tyr Ile Cys Asp Glu  
 25 30 35  
 tgc att gag ctg tgc aac gag att att gaa gaa gaa ctc ggt caa gct 259  
 Cys Ile Glu Leu Cys Asn Glu Ile Ile Glu Glu Glu Leu Gly Gln Ala  
 40 45 50  
 caa cac gac gag cag gag cgc aac gag ctc ccc aag ccg tcg gag att 307  
 Gln His Asp Glu Gln Glu Arg Asn Glu Leu Pro Lys Pro Ser Glu Ile  
 55 60 65  
 tca gcc ttc ctt gat act tat gtc atc ggg cag gac cca gca aaa cgt 355  
 Ser Ala Phe Leu Asp Thr Tyr Val Ile Gly Gln Asp Pro Ala Lys Arg  
 70 75 80 85  
 atc ctg tcg gtt gcg gtg tac aac cat tac aag cgt ctc cgc gca tcg 403  
 Ile Leu Ser Val Ala Val Tyr Asn His Tyr Lys Arg Leu Arg Ala Ser  
 90 95 100  
 gaa acc atc ggt cgt cgc agg aat gac gag cct gaa acc gaa ctg gtt 451  
 Glu Thr Ile Gly Arg Arg Arg Asn Asp Glu Pro Glu Thr Glu Leu Val  
 105 110 115

aag tcc aat att ttg atg ctc ggc ccc act ggc tcc ggc aag act ttc	499
Lys Ser Asn Ile Leu Met Leu Gly Pro Thr Gly Ser Gly Lys Thr Phe	
120 125 130	
ctt gcc cag act ttg gca aag ctg ctg gat gtt cct ttt gct atc gcg	547
Leu Ala Gln Thr Leu Ala Lys Leu Leu Asp Val Pro Phe Ala Ile Ala	
135 140 145	
gat gcc acc tca ctg acc gag gct ggt tat gtg ggc gag gat gtg gaa	595
Asp Ala Thr Ser Leu Thr Glu Ala Gly Tyr Val Gly Glu Asp Val Glu	
150 155 160 165	
aac atc ttg ctc aag ctg ctt cag gct gct gat ttt gat gtg gaa cgt	643
Asn Ile Leu Leu Lys Leu Leu Gln Ala Ala Asp Phe Asp Val Glu Arg	
170 175 180	
gca cag cgc ggc atc att tac atc gat gaa gtg gac aag att tcc cgc	691
Ala Gln Arg Gly Ile Ile Tyr Ile Asp Glu Val Asp Lys Ile Ser Arg	
185 190 195	
aag tct gaa aac cca tcg atc act cgc gat gtt tcc ggt gaa ggc gtg	739
Lys Ser Glu Asn Pro Ser Ile Thr Arg Asp Val Ser Gly Glu Gly Val	
200 205 210	
cag cag gca ctg ctg aaa att ttg gaa ggc act gtc gcc gca atc cca	787
Gln Gln Ala Leu Leu Lys Ile Leu Glu Gly Thr Val Ala Ala Ile Pro	
215 220 225	
ccg cag gga gga cgc aag cac ccc aac cag gat ttc atc cag ctg gat	835
Pro Gln Gly Gly Arg Lys His Pro Asn Gln Asp Phe Ile Gln Leu Asp	
230 235 240 245	
acc acc aac att ttg ttc atc gtt gct ggt gcg ttc tct ggt ctg gag	883
Thr Thr Asn Ile Leu Phe Ile Val Ala Gly Ala Phe Ser Gly Leu Glu	
250 255 260	
aag gtc atc gcg gac cgc aat ggc aag aaa ggc ttg ggc ttc ggt gtg	931
Lys Val Ile Ala Asp Arg Asn Gly Lys Lys Gly Leu Gly Phe Gly Val	
265 270 275	
gag gtc tct tcc aag aag gaa gaa gcc aac att gtg gat atc ttc aag	979
Glu Val Ser Ser Lys Lys Glu Glu Ala Asn Ile Val Asp Ile Phe Lys	
280 285 290	
gat gtc ctc cct gag gac ctg gtg aag ttt ggt ctc atc cca gaa ttc	1027
Asp Val Leu Pro Glu Asp Leu Val Lys Phe Gly Leu Ile Pro Glu Phe	
295 300 305	
att ggg cgt ctg cca gtc gtt gcc acc gta tcc aac ctg gat cag aaa	1075
Ile Gly Arg Leu Pro Val Val Ala Thr Val Ser Asn Leu Asp Gln Lys	
310 315 320 325	
tct ctg gtc aag gtt ctc acg gag cct cgt aac tca ttg gtg aag cag	1123
Ser Leu Val Lys Val Leu Thr Glu Pro Arg Asn Ser Leu Val Lys Gln	
330 335 340	
tat cga cgt ctg ttt gaa atg gat gac gct gtg ttg acc ttt act gat	1171
Tyr Arg Arg Leu Phe Glu Met Asp Asp Ala Val Leu Thr Phe Thr Asp	
345 350 355	
gat gct ttg gag gag atc gct aat cag gca ctc gag cgc aaa act ggc	1219

Asp Ala Leu Glu Glu Ile Ala Asn Gln Ala Leu Glu Arg Lys Thr Gly  
 360 365 370

gcc cgt ggc ctg cgc gcg atc atg gaa gag atc ctg gtt ccg atc atg 1267  
 Ala Arg Gly Leu Arg Ala Ile Met Glu Glu Ile Leu Val Pro Ile Met  
 375 380 385

tat gac ctg cca gac cgt aaa gac gtt ggc gaa gtc atc atc aac ggt 1315  
 Tyr Asp Leu Pro Asp Arg Lys Asp Val Gly Glu Val Ile Ile Asn Gly  
 390 395 400 405

gcc gtt gcc cgt ggc gaa gcc gaa cca gag atg ttg gaa gct gtc gca 1363  
 Ala Val Ala Arg Gly Glu Ala Glu Pro Glu Met Leu Glu Ala Val Ala  
 410 415 420

gaa gaa aag acc gcg tagttggcag gagttatcac cgg 1401  
 Glu Glu Lys Thr Ala  
 425

<210> 70  
 <211> 426  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 70  
 Met Ala Arg Met Gln Glu Ser Ala Asp Leu Leu Lys Cys Ser Phe Cys  
 1 5 10 15

Gly Lys Ser Gln Lys Gln Val Lys Lys Leu Ile Ala Gly Gly Ala Val  
 20 25 30

Tyr Ile Cys Asp Glu Cys Ile Glu Leu Cys Asn Glu Ile Ile Glu Glu  
 35 40 45

Glu Leu Gly Gln Ala Gln His Asp Glu Gln Glu Arg Asn Glu Leu Pro  
 50 55 60

Lys Pro Ser Glu Ile Ser Ala Phe Leu Asp Thr Tyr Val Ile Gly Gln  
 65 70 75 80

Asp Pro Ala Lys Arg Ile Leu Ser Val Ala Val Tyr Asn His Tyr Lys  
 85 90 95

Arg Leu Arg Ala Ser Glu Thr Ile Gly Arg Arg Arg Asn Asp Glu Pro  
 100 105 110

Glu Thr Glu Leu Val Lys Ser Asn Ile Leu Met Leu Gly Pro Thr Gly  
 115 120 125

Ser Gly Lys Thr Phe Leu Ala Gln Thr Leu Ala Lys Leu Leu Asp Val  
 130 135 140

Pro Phe Ala Ile Ala Asp Ala Thr Ser Leu Thr Glu Ala Gly Tyr Val  
 145 150 155 160

Gly Glu Asp Val Glu Asn Ile Leu Leu Lys Leu Leu Gln Ala Ala Asp  
 165 170 175

Phe Asp Val Glu Arg Ala Gln Arg Gly Ile Ile Tyr Ile Asp Glu Val  
 180 185 190

Asp Lys Ile Ser Arg Lys Ser Glu Asn Pro Ser Ile Thr Arg Asp Val  
 195 200 205  
 Ser Gly Glu Gly Val Gln Gln Ala Leu Leu Lys Ile Leu Glu Gly Thr  
 210 215 220  
 Val Ala Ala Ile Pro Pro Gln Gly Gly Arg Lys His Pro Asn Gln Asp  
 225 230 235 240  
 Phe Ile Gln Leu Asp Thr Thr Asn Ile Leu Phe Ile Val Ala Gly Ala  
 245 250 255  
 Phe Ser Gly Leu Glu Lys Val Ile Ala Asp Arg Asn Gly Lys Lys Gly  
 260 265 270  
 Leu Gly Phe Gly Val Glu Val Ser Ser Lys Lys Glu Glu Ala Asn Ile  
 275 280 285  
 Val Asp Ile Phe Lys Asp Val Leu Pro Glu Asp Leu Val Lys Phe Gly  
 290 295 300  
 Leu Ile Pro Glu Phe Ile Gly Arg Leu Pro Val Val Ala Thr Val Ser  
 305 310 315 320  
 Asn Leu Asp Gln Lys Ser Leu Val Lys Val Leu Thr Glu Pro Arg Asn  
 325 330 335  
 Ser Leu Val Lys Gln Tyr Arg Arg Leu Phe Glu Met Asp Asp Ala Val  
 340 345 350  
 Leu Thr Phe Thr Asp Asp Ala Leu Glu Glu Ile Ala Asn Gln Ala Leu  
 355 360 365  
 Glu Arg Lys Thr Gly Ala Arg Gly Leu Arg Ala Ile Met Glu Glu Ile  
 370 375 380  
 Leu Val Pro Ile Met Tyr Asp Leu Pro Asp Arg Lys Asp Val Gly Glu  
 385 390 395 400  
 Val Ile Ile Asn Gly Ala Val Ala Arg Gly Glu Ala Glu Pro Glu Met  
 405 410 415  
 Leu Glu Ala Val Ala Glu Glu Lys Thr Ala  
 420 425

&lt;210&gt; 71

&lt;211&gt; 1401

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (101)..(1378)

&lt;223&gt; FRXA01120

&lt;400&gt; 71

acaggtaaag cgctaagatg gaacaacca ttgccaatat tgttggttag agttgtacgc 60

agtaaattctt ttcaatcgtg gaagcgggtc tcacagtcta atg gca cgt atg cag 115



																Met	Ala	Arg	Met	Gln	
																1				5	
gaa	agc	gcc	gat	ctg	ctc	aaa	tgt	tcc	ttc	tgc	gga	aag	agc	caa	aag	163					
Glu	Ser	Ala	Asp	Leu	Leu	Lys	Cys	Ser	Phe	Cys	Gly	Lys	Ser	Gln	Lys						
				10					15					20							
cag	gta	aaa	aaa	ctc	atc	gcg	ggt	ggc	gcc	gta	tat	atc	tgt	gat	gag	211					
Gln	Val	Lys	Lys	Leu	Ile	Ala	Gly	Gly	Ala	Val	Tyr	Ile	Cys	Asp	Glu						
				25					30					35							
tgc	att	gag	ctg	tgc	aac	gag	att	att	gaa	gaa	gaa	ctc	ggt	caa	gct	259					
Cys	Ile	Glu	Leu	Cys	Asn	Glu	Ile	Ile	Glu	Glu	Glu	Leu	Gly	Gln	Ala						
				40					45					50							
caa	cac	gac	gag	cag	gag	cgc	aac	gag	ctc	ccc	aag	ccg	tcg	gag	att	307					
Gln	His	Asp	Glu	Gln	Glu	Arg	Asn	Glu	Leu	Pro	Lys	Pro	Ser	Glu	Ile						
				55					60					65							
tca	gcc	ttc	ctt	gat	act	tat	gtc	atc	ggg	cag	gac	cca	gca	aaa	cgt	355					
Ser	Ala	Phe	Leu	Asp	Thr	Tyr	Val	Ile	Gly	Gln	Asp	Pro	Ala	Lys	Arg						
				70					75					80							
atc	ctg	tcg	gtt	gcg	gtg	tac	aac	cat	tac	aag	cgt	ctc	cgc	gca	tcg	403					
Ile	Leu	Ser	Val	Ala	Val	Tyr	Asn	His	Tyr	Lys	Arg	Leu	Arg	Ala	Ser						
				90					95					100							
gaa	acc	atc	ggt	cgt	cgc	agg	aat	gac	gag	cct	gaa	acc	gaa	ctg	gtt	451					
Glu	Thr	Ile	Gly	Arg	Arg	Arg	Asn	Asp	Glu	Pro	Glu	Thr	Glu	Leu	Val						
				105					110					115							
aag	tcc	aat	att	ttg	atg	ctc	ggc	ccc	act	ggc	tcc	ggc	aag	act	ttc	499					
Lys	Ser	Asn	Ile	Leu	Met	Leu	Gly	Pro	Thr	Gly	Ser	Gly	Lys	Thr	Phe						
				120					125					130							
ctt	gcc	cag	act	ttg	gca	aag	ctg	ctg	gat	gtt	cct	ttt	gct	atc	gcg	547					
Leu	Ala	Gln	Thr	Leu	Ala	Lys	Leu	Leu	Asp	Val	Pro	Phe	Ala	Ile	Ala						
				135					140					145							
gat	gcc	acc	tca	ctg	acc	gag	gct	ggt	tat	gtg	ggc	gag	gat	gtg	gaa	595					
Asp	Ala	Thr	Ser	Leu	Thr	Glu	Ala	Gly	Tyr	Val	Gly	Glu	Asp	Val	Glu						
				150					155					160							
aac	atc	ttg	ctc	aag	ctg	ctt	cag	gct	gct	gat	ttt	gat	gtg	gaa	cgt	643					
Asn	Ile	Leu	Leu	Lys	Leu	Leu	Gln	Ala	Ala	Asp	Phe	Asp	Val	Glu	Arg						
				170					175					180							
gca	cag	cgc	ggc	atc	att	tac	atc	gat	gaa	gtg	gac	aag	att	tcc	cgc	691					
Ala	Gln	Arg	Gly	Ile	Ile	Tyr	Ile	Asp	Glu	Val	Asp	Lys	Ile	Ser	Arg						
				185					190					195							
aag	tct	gaa	aac	cca	tcg	atc	act	cgc	gat	gtt	tcc	ggt	gaa	ggc	gtg	739					
Lys	Ser	Glu	Asn	Pro	Ser	Ile	Thr	Arg	Asp	Val	Ser	Gly	Glu	Gly	Val						
				200					205					210							
cag	cag	gca	ctg	ctg	aaa	att	ttg	gaa	ggc	act	gtc	gcc	gca	atc	cca	787					
Gln	Gln	Ala	Leu	Leu	Lys	Ile	Leu	Glu	Gly	Thr	Val	Ala	Ala	Ile	Pro						
				215					220					225							
ccg	cag	gga	gga	cgc	aag	cac	ccc	aac	cag	gat	ttc	atc	cag	ctg	gat	835					
Pro	Gln	Gly	Gly	Arg	Lys	His	Pro	Asn	Gln	Asp	Phe	Ile	Gln	Leu	Asp						

230	235	240	245	
acc acc aac att ttg ttc atc gtt gct ggt gcg ttc tct ggt ctg gag	883			
Thr Thr Asn Ile Leu Phe Ile Val Ala Gly Ala Phe Ser Gly Leu Glu				
250	255	260		
aag gtc atc gcg gac cgc aat ggc aag aaa ggc ttg ggc ttc ggt gtg	931			
Lys Val Ile Ala Asp Arg Asn Gly Lys Lys Gly Leu Gly Phe Gly Val				
265	270	275		
gag gtc tct tcc aag aag gaa gaa gcc aac att gtg gat atc ttc aag	979			
Glu Val Ser Ser Lys Lys Glu Glu Ala Asn Ile Val Asp Ile Phe Lys				
280	285	290		
gat gtc ctc cct gag gac ctg gtg aag ttt ggt ctc atc cca gaa ttc	1027			
Asp Val Leu Pro Glu Asp Leu Val Lys Phe Gly Leu Ile Pro Glu Phe				
295	300	305		
att ggg cgt ctg cca gtc gtt gcc acc gta tcc aac ctg gat cag aaa	1075			
Ile Gly Arg Leu Pro Val Val Ala Thr Val Ser Asn Leu Asp Gln Lys				
310	315	320	325	
tct ctg gtc aag gtt ctc acg gag cct cgt aac tca ttg gtg aag cag	1123			
Ser Leu Val Lys Val Leu Thr Glu Pro Arg Asn Ser Leu Val Lys Gln				
330	335	340		
tat cga cgt ctg ttt gaa atg gat gac gct gtg ttg acc ttt act gat	1171			
Tyr Arg Arg Leu Phe Glu Met Asp Asp Ala Val Leu Thr Phe Thr Asp				
345	350	355		
gat gct ttg gag gag atc gct aat cag gca ctc gag cgc aaa act ggc	1219			
Asp Ala Leu Glu Glu Ile Ala Asn Gln Ala Leu Glu Arg Lys Thr Gly				
360	365	370		
gcc cgt ggc ctg cgc gcg atc atg gaa gag atc ctg gtt ccg atc atg	1267			
Ala Arg Gly Leu Arg Ala Ile Met Glu Glu Ile Leu Val Pro Ile Met				
375	380	385		
tat gac ctc cca gac cgt aaa gac gtt ggc gaa gtc atc atc aac ggt	1315			
Tyr Asp Leu Pro Asp Arg Lys Asp Val Gly Glu Val Ile Ile Asn Gly				
390	395	400	405	
gcc gtt gcc cgt ggc gaa gcc gaa cca gag atg ttg gaa gct gtc gca	1363			
Ala Val Ala Arg Gly Glu Ala Glu Pro Glu Met Leu Glu Ala Val Ala				
410	415	420		
gaa gaa aag acc gcg tagttggcag gagttatcac cgg	1401			
Glu Glu Lys Thr Ala				
425				

&lt;210&gt; 72

&lt;211&gt; 426

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 72

Met	Ala	Arg	Met	Gln	Glu	Ser	Ala	Asp	Leu	Leu	Lys	Cys	Ser	Phe	Cys
1				5					10					15	

Gly Lys Ser Gln Lys Gln Val Lys Lys Leu Ile Ala Gly Gly Ala Val

20					25					30					
Tyr	Ile	Cys	Asp	Glu	Cys	Ile	Glu	Leu	Cys	Asn	Glu	Ile	Ile	Glu	Glu
		35					40					45			
Glu	Leu	Gly	Gln	Ala	Gln	His	Asp	Glu	Gln	Glu	Arg	Asn	Glu	Leu	Pro
	50					55					60				
Lys	Pro	Ser	Glu	Ile	Ser	Ala	Phe	Leu	Asp	Thr	Tyr	Val	Ile	Gly	Gln
65					70				75						80
Asp	Pro	Ala	Lys	Arg	Ile	Leu	Ser	Val	Ala	Val	Tyr	Asn	His	Tyr	Lys
				85					90					95	
Arg	Leu	Arg	Ala	Ser	Glu	Thr	Ile	Gly	Arg	Arg	Arg	Asn	Asp	Glu	Pro
			100					105					110		
Glu	Thr	Glu	Leu	Val	Lys	Ser	Asn	Ile	Leu	Met	Leu	Gly	Pro	Thr	Gly
	115						120					125			
Ser	Gly	Lys	Thr	Phe	Leu	Ala	Gln	Thr	Leu	Ala	Lys	Leu	Leu	Asp	Val
130						135					140				
Pro	Phe	Ala	Ile	Ala	Asp	Ala	Thr	Ser	Leu	Thr	Glu	Ala	Gly	Tyr	Val
145					150					155					160
Gly	Glu	Asp	Val	Glu	Asn	Ile	Leu	Leu	Lys	Leu	Leu	Gln	Ala	Ala	Asp
				165				170					175		
Phe	Asp	Val	Glu	Arg	Ala	Gln	Arg	Gly	Ile	Ile	Tyr	Ile	Asp	Glu	Val
		180						185					190		
Asp	Lys	Ile	Ser	Arg	Lys	Ser	Glu	Asn	Pro	Ser	Ile	Thr	Arg	Asp	Val
	195						200					205			
Ser	Gly	Glu	Gly	Val	Gln	Gln	Ala	Leu	Leu	Lys	Ile	Leu	Glu	Gly	Thr
210						215					220				
Val	Ala	Ala	Ile	Pro	Pro	Gln	Gly	Gly	Arg	Lys	His	Pro	Asn	Gln	Asp
225					230					235					240
Phe	Ile	Gln	Leu	Asp	Thr	Thr	Asn	Ile	Leu	Phe	Ile	Val	Ala	Gly	Ala
				245					250					255	
Phe	Ser	Gly	Leu	Glu	Lys	Val	Ile	Ala	Asp	Arg	Asn	Gly	Lys	Lys	Gly
			260					265					270		
Leu	Gly	Phe	Gly	Val	Glu	Val	Ser	Ser	Lys	Lys	Glu	Glu	Ala	Asn	Ile
		275					280					285			
Val	Asp	Ile	Phe	Lys	Asp	Val	Leu	Pro	Glu	Asp	Leu	Val	Lys	Phe	Gly
	290					295					300				
Leu	Ile	Pro	Glu	Phe	Ile	Gly	Arg	Leu	Pro	Val	Val	Ala	Thr	Val	Ser
305						310				315					320
Asn	Leu	Asp	Gln	Lys	Ser	Leu	Val	Lys	Val	Leu	Thr	Glu	Pro	Arg	Asn
				325					330					335	
Ser	Leu	Val	Lys	Gln	Tyr	Arg	Arg	Leu	Phe	Glu	Met	Asp	Asp	Ala	Val
			340					345					350		

Leu Thr Phe Thr Asp Asp Ala Leu Glu Glu Ile Ala Asn Gln Ala Leu  
 355 360 365  
 Glu Arg Lys Thr Gly Ala Arg Gly Leu Arg Ala Ile Met Glu Glu Ile  
 370 375 380  
 Leu Val Pro Ile Met Tyr Asp Leu Pro Asp Arg Lys Asp Val Gly Glu  
 385 390 395 400  
 Val Ile Ile Asn Gly Ala Val Ala Arg Gly Glu Ala Glu Pro Glu Met  
 405 410 415  
 Leu Glu Ala Val Ala Glu Glu Lys Thr Ala  
 420 425

<210> 73  
 <211> 1065  
 <212> DNA  
 <213> *Corynebacterium glutamicum*

<220>  
 <221> CDS  
 <222> (101)..(1042)  
 <223> RXA00744

<400> 73  
 tctgagtcg tagaagtatt acccagtgac ttaagtttct tagatttttt tgagcaacag 60  
 cgaccagcca cgtagtggtg gtcgagtaga ggatagctac atg ggg aac tgg gca 115  
 Met Gly Asn Trp Ala  
 1 5  
 gag att act gat gaa att tct aag att tac caa gat aat cag tac aag 163  
 Glu Ile Thr Asp Glu Ile Ser Lys Ile Tyr Gln Asp Asn Gln Tyr Lys  
 10 15 20  
 att aga caa ata aat gat gtt gac gca gta agc gat aaa cgt aga gaa 211  
 Ile Arg Gln Ile Asn Asp Val Asp Ala Val Ser Asp Lys Arg Arg Glu  
 25 30 35  
 gcg cta caa gca ctg ttt gaa cat act ggt cga aat gta atc gtc tat 259  
 Ala Leu Gln Ala Leu Phe Glu His Thr Gly Arg Asn Val Ile Val Tyr  
 40 45 50  
 tat tca gcg tgg tta gaa aat ggt cga cga ttt tcc ggg caa tct acg 307  
 Tyr Ser Ala Trp Leu Glu Asn Gly Arg Arg Phe Ser Gly Gln Ser Thr  
 55 60 65  
 gat ttt tcg gta aat gat act gat aaa aac agt ttt atg act gcg ctc 355  
 Asp Phe Ser Val Asn Asp Thr Asp Lys Asn Ser Phe Met Thr Ala Leu  
 70 75 80 85  
 cat aag ttg gat cag agt aaa ggt ctc gat ctt atc ctc cac act ccg 403  
 His Lys Leu Asp Gln Ser Lys Gly Leu Asp Leu Ile Leu His Thr Pro  
 90 95 100  
 ggt gga gat gtt gct gcg aca gag tcg tta gta gat tac att cac gca 451  
 Gly Gly Asp Val Ala Ala Thr Glu Ser Leu Val Asp Tyr Ile His Ala  
 105 110 115

```

ctc ttt ggt caa gat ttc aga gtc att gtc ccc caa ctc gca atg tca 499
Leu Phe Gly Gln Asp Phe Arg Val Ile Val Pro Gln Leu Ala Met Ser
      120                      125                      130

gca gga aca atg atc gca ctt tcg tcc aaa gag att gtt atg ggg aag 547
Ala Gly Thr Met Ile Ala Leu Ser Ser Lys Glu Ile Val Met Gly Lys
      135                      140                      145

cat tct agt ctt ggc ccc att gat cct cag ttt aac ggc cta ccg gca 595
His Ser Ser Leu Gly Pro Ile Asp Pro Gln Phe Asn Gly Leu Pro Ala
      150                      155                      160                      165

cac ggg tta ttg gaa gaa ttt gag caa gcg aag aaa gag gtc tct gag 643
His Gly Leu Leu Glu Glu Phe Glu Gln Ala Lys Lys Glu Val Ser Glu
      170                      175                      180

aat ccg cag act gct cat ata tgg cag gtg atc ttg aat aaa tac aac 691
Asn Pro Gln Thr Ala His Ile Trp Gln Val Ile Leu Asn Lys Tyr Asn
      185                      190                      195

ccc acg atg ttg ggt gaa gct aaa aaa gct att cag tgg tcc aac tcg 739
Pro Thr Met Leu Gly Glu Ala Lys Lys Ala Ile Gln Trp Ser Asn Ser
      200                      205                      210

atg gtt aag cag tgg ctt gaa aag ggt atg ttt tta gac gag cct gac 787
Met Val Lys Gln Trp Leu Glu Lys Gly Met Phe Leu Asp Glu Pro Asp
      215                      220                      225

aaa gaa gaa aaa gcc act cgc gct atc aaa gag ctc gct gat cat tcc 835
Lys Glu Glu Lys Ala Thr Arg Ala Ile Lys Glu Leu Ala Asp His Ser
      230                      235                      240                      245

gtt act ctt gcg cat aat cga cac att tcg gtc agt aaa gca ctt gag 883
Val Thr Leu Ala His Asn Arg His Ile Ser Val Ser Lys Ala Leu Glu
      250                      255                      260

ctg gga ttg aat atc aaa gaa ctt gag agc gat cca aag ctt caa gat 931
Leu Gly Leu Asn Ile Lys Glu Leu Glu Ser Asp Pro Lys Leu Gln Asp
      265                      270                      275

tta gtt ctt act ctt cac cac ctg tcc gtt att gct gcg caa cga gga 979
Leu Val Leu Thr Leu His His Leu Ser Val Ile Ala Ala Gln Arg Gly
      280                      285                      290

cca tta att aag ttt gtc gtc aat cat gac aac cgt ggc act ttt ctg 1027
Pro Leu Ile Lys Phe Val Val Asn His Asp Asn Arg Gly Thr Phe Leu
      295                      300                      305

cag ggg cat gaa aac taattaagt atgcaatagt cta 1065
Gln Gly His Glu Asn
310

```

&lt;210&gt; 74

&lt;211&gt; 314

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 74

Met Gly Asn Trp Ala Glu Ile Thr Asp Glu Ile Ser Lys Ile Tyr Gln

1	5	10	15
Asp Asn Gln Tyr	Lys Ile Arg Gln Ile	Asn Asp Val Asp	Ala Val Ser
20	25	30	
Asp Lys Arg Arg	Glu Ala Leu Gln Ala	Leu Phe Glu His	Thr Gly Arg
35	40	45	
Asn Val Ile Val	Tyr Tyr Ser Ala	Trp Leu Glu Asn	Gly Arg Arg Phe
50	55	60	
Ser Gly Gln Ser	Thr Asp Phe Ser	Val Asn Asp Thr	Asp Lys Asn Ser
65	70	75	80
Phe Met Thr Ala	Leu His Lys Leu	Asp Gln Ser Lys	Gly Leu Asp Leu
85	90	95	
Ile Leu His Thr	Pro Gly Gly Asp	Val Ala Ala Thr	Glu Ser Leu Val
100	105	110	
Asp Tyr Ile His	Ala Leu Phe Gly	Gln Asp Phe Arg	Val Ile Val Pro
115	120	125	
Gln Leu Ala Met	Ser Ala Gly Thr	Met Ile Ala Leu	Ser Ser Lys Glu
130	135	140	
Ile Val Met Gly	Lys His Ser Ser	Leu Gly Pro Ile	Asp Pro Gln Phe
145	150	155	160
Asn Gly Leu Pro	Ala His Gly Leu	Leu Glu Glu Phe	Glu Gln Ala Lys
165	170	175	
Lys Glu Val Ser	Glu Asn Pro Gln	Thr Ala His Ile	Trp Gln Val Ile
180	185	190	
Leu Asn Lys Tyr	Asn Pro Thr Met	Leu Gly Glu Ala	Lys Lys Ala Ile
195	200	205	
Gln Trp Ser Asn	Ser Met Val Lys	Gln Trp Leu Glu	Lys Gly Met Phe
210	215	220	
Leu Asp Glu Pro	Asp Lys Glu Glu	Lys Ala Thr Arg	Ala Ile Lys Glu
225	230	235	240
Leu Ala Asp His	Ser Val Thr Leu	Ala His Asn Arg	His Ile Ser Val
245	250	255	
Ser Lys Ala Leu	Glu Leu Gly Leu	Asn Ile Lys Glu	Leu Glu Ser Asp
260	265	270	
Pro Lys Leu Gln	Asp Leu Val Leu	Thr Leu His His	Leu Ser Val Ile
275	280	285	
Ala Ala Gln Arg	Gly Pro Leu Ile	Lys Phe Val Val	Asn His Asp Asn
290	295	300	
Arg Gly Thr Phe	Leu Gln Gly His	Glu Asn	
305	310		

&lt;210&gt; 75

```
<220>
<221> CDS
<222> (101)..(934)
<223> RXA00844
```

ggtgacggcc	aaaggcgcgc	gcattgcgcg	ccccaaaatc	gacgtcattg	acagcatctt	60
gtgacaacat	tttgtagagc	aaccatctag	actgttcttt	atg tct tct gcg tca		115
				Met Ser Ser Ala Ser		
				1	5	
ttt acc acc aaa gca ctg tcc gta ctc gca gct tta acg gct gcg tct						163
Phe Thr Thr Lys Ala Leu Ser Val Leu Ala Ala Leu Thr Ala Ala Ser						
	10		15		20	
gcc ccc tta gtg gcg gcg tca cct gca cat gct ttg gca aat gct cgc						211
Ala Pro Leu Val Ala Ala Ser Pro Ala His Ala Leu Ala Asn Ala Arg						
	25		30		35	
aac gtt acg ggt tca agc acc act tca gat tca att gtt cgt ctg cac						259
Asn Val Thr Gly Ser Ser Thr Thr Ser Asp Ser Ile Val Arg Leu His						
	40		45		50	
atc ggt aac act gca tgt aca gga acc atg atc acc cca acg tgg gcg						307
Ile Gly Asn Thr Ala Cys Thr Gly Thr Met Ile Thr Pro Thr Trp Ala						
	55		60		65	
atc acc gcc cgc cac tgt atc cct gag ggc ggt att gcc ggt gca gct						355
Ile Thr Ala Arg His Cys Ile Pro Glu Gly Gly Ile Ala Gly Ala Ala						
	70		75		80	85
att ggt tca agc act ttg agc caa ttt cag cag gtg tcc caa gcg atc						403
Ile Gly Ser Ser Thr Leu Ser Gln Phe Gln Gln Val Ser Gln Ala Ile						
	90		95		100	
ttg cac cct act gcg gac tta gct ctc gtt gag ctt ccc aat cag gca						451
Leu His Pro Thr Ala Asp Leu Ala Leu Val Glu Leu Pro Asn Gln Ala						
	105		110		115	
agt tcc aac acg gtt gat ctc tac ggt gca cac gtg cag cct ggt gaa						499
Ser Ser Asn Thr Val Asp Leu Tyr Gly Ala His Val Gln Pro Gly Glu						
	120		125		130	
aat ggt caa gca gcc ggc tgg ggt ggg tac tct gcc ttt ggc caa aat						547
Asn Gly Gln Ala Ala Gly Trp Gly Gly Tyr Ser Ala Phe Gly Gln Asn						
	135		140		145	
ggt gca cag caa gcc gat gtg caa att caa cgc agg gta gtc aat gtg						595
Val Ala Gln Gln Ala Asp Val Gln Ile Gln Arg Arg Val Val Asn Val						
	150		155		160	165
cca agc ccc gac cgc acc gct gtg ctg ctt gaa ggc act gtt tct aac						643
Pro Ser Pro Asp Arg Thr Ala Val Leu Leu Glu Gly Thr Val Ser Asn						
	170		175		180	
ggt cgt ctc gta cca gcc gat tcc gcc gga cct ttg tac atc aat ggt						691

Gly Arg Leu Val Pro Gly Asp Ser Gly Gly Pro Leu Tyr Ile Asn Gly  
 185 190 195

caa ctg gct ggt gtg ctc agc atg tcc act gac gta gaa aac gat gca 739  
 Gln Leu Ala Gly Val Leu Ser Met Ser Thr Asp Val Glu Asn Asp Ala  
 200 205 210

cta gac ggc acc gtc ggc tgg tac atc ccc gtt gct gaa cac gcc gag 787  
 Leu Asp Gly Thr Val Gly Trp Tyr Ile Pro Val Ala Glu His Ala Glu  
 215 220 225

tgg atc gcc tac tac acc ggc aag cac att gcc ccc att gct ggt gcg 835  
 Trp Ile Ala Tyr Tyr Thr Gly Lys His Ile Ala Pro Ile Ala Gly Ala  
 230 235 240 245

ccc gca gaa ctt gtt gac gcc acc gcc aac ccc acc ttc atc cct gct 883  
 Pro Ala Glu Leu Val Asp Ala Thr Ala Asn Pro Thr Phe Ile Pro Ala  
 250 255 260

cca cag cct ttc acc ggt tca tcc atc ggt ggt tgg gcg ctg ggc agc 931  
 Pro Gln Pro Phe Thr Gly Ser Ser Ile Gly Gly Trp Ala Leu Gly Ser  
 265 270 275

tcc tagaatatgc tgatctccct gct 957  
 Ser

<210> 76  
 <211> 278  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 76  
 Met Ser Ser Ala Ser Phe Thr Thr Lys Ala Leu Ser Val Leu Ala Ala  
 1 5 10 15

Leu Thr Ala Ala Ser Ala Pro Leu Val Ala Ala Ser Pro Ala His Ala  
 20 25 30

Leu Ala Asn Ala Arg Asn Val Thr Gly Ser Ser Thr Thr Ser Asp Ser  
 35 40 45

Ile Val Arg Leu His Ile Gly Asn Thr Ala Cys Thr Gly Thr Met Ile  
 50 55 60

Thr Pro Thr Trp Ala Ile Thr Ala Arg His Cys Ile Pro Glu Gly Gly  
 65 70 75 80

Ile Ala Gly Ala Ala Ile Gly Ser Ser Thr Leu Ser Gln Phe Gln Gln  
 85 90 95

Val Ser Gln Ala Ile Leu His Pro Thr Ala Asp Leu Ala Leu Val Glu  
 100 105 110

Leu Pro Asn Gln Ala Ser Ser Asn Thr Val Asp Leu Tyr Gly Ala His  
 115 120 125

Val Gln Pro Gly Glu Asn Gly Gln Ala Ala Gly Trp Gly Gly Tyr Ser  
 130 135 140



Ala Phe Gly Gln Asn Val Ala Gln Gln Ala Asp Val Gln Ile Gln Arg  
 145 150 155 160

Arg Val Val Asn Val Pro Ser Pro Asp Arg Thr Ala Val Leu Leu Glu  
 165 170 175

Gly Thr Val Ser Asn Gly Arg Leu Val Pro Gly Asp Ser Gly Gly Pro  
 180 185 190

Leu Tyr Ile Asn Gly Gln Leu Ala Gly Val Leu Ser Met Ser Thr Asp  
 195 200 205

Val Glu Asn Asp Ala Leu Asp Gly Thr Val Gly Trp Tyr Ile Pro Val  
 210 215 220

Ala Glu His Ala Glu Trp Ile Ala Tyr Tyr Thr Gly Lys His Ile Ala  
 225 230 235 240

Pro Ile Ala Gly Ala Pro Ala Glu Leu Val Asp Ala Thr Ala Asn Pro  
 245 250 255

Thr Phe Ile Pro Ala Pro Gln Pro Phe Thr Gly Ser Ser Ile Gly Gly  
 260 265 270

Trp Ala Leu Gly Ser Ser  
 275

<210> 77  
 <211> 958  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(958)  
 <223> RXA01151

<400> 77  
 ccaggcggttt ggctgatggc gtggagcctc ctcgtccgca gcgtaaggca cgtcgataaa 60

aagacctagt tggagggcgt aaaggggttag agtggtgacc atg agt tca cca act 115  
 Met Ser Ser Pro Thr  
 1 5

gat tct tcg ccc tct aat tct ttt agc gac ttc aac cgg gag gaa cag 163  
 Asp Ser Ser Pro Ser Asn Ser Phe Ser Asp Phe Asn Arg Glu Glu Gln  
 10 15 20

tcc cgg tta tot gat gag gtg cgc cag ctc aag cgc acc aac tct gat 211  
 Ser Arg Leu Ser Asp Glu Val Arg Gln Leu Lys Arg Thr Asn Ser Asp  
 25 30 35

ctt ggg gca cgt aat gcc aag ctc gcg gag atg ctg aag tcg tct cgg 259  
 Leu Gly Ala Arg Asn Ala Lys Leu Ala Glu Met Leu Lys Ser Ser Arg  
 40 45 50

gat aaa ttg tct gtg ctg ttt tct cag ttg gag gat atg gct cag ccg 307  
 Asp Lys Leu Ser Val Leu Phe Ser Gln Leu Glu Asp Met Ala Gln Pro  
 55 60 65

cca tcg gtg tat ggc act ttc ttg gaa acc gcg aaa gac ggt tct aat	355
Pro Ser Val Tyr Gly Thr Phe Leu Glu Thr Ala Lys Asp Gly Ser Asn	
70 75 80 85	
gcg gag atc ttt gct ggt gga cgt cgc atg cgt gtg gct gtt tct cct	403
Ala Glu Ile Phe Ala Gly Gly Arg Arg Met Arg Val Ala Val Ser Pro	
90 95 100	
atg ctg tgt gcc gcg gat ttg atg ccg ggt gtg cag gtt cgt ttg ggt	451
Met Leu Cys Ala Ala Asp Leu Met Pro Gly Val Gln Val Arg Leu Gly	
105 110 115	
gaa ggc aat caa gtt ctt gag gcc tgt gat ttt gaa caa acc ggt gaa	499
Glu Gly Asn Gln Val Leu Glu Ala Cys Asp Phe Glu Gln Thr Gly Glu	
120 125 130	
tta gcc acg ttg atg gaa atg att ggc cgg gat cgt gct ttg gtt tca	547
Leu Ala Thr Leu Met Glu Met Ile Gly Arg Asp Arg Ala Leu Val Ser	
135 140 145	
gat cgc tcg ggg gag gag cgc gtc gtc aag ctt gct ggt ccg ttg atg	595
Asp Arg Ser Gly Glu Glu Arg Val Val Lys Leu Ala Gly Pro Leu Met	
150 155 160 165	
gat cgc acc gca aag ctg ccg cgc ccc ggt gac acc ctg ctt gtt gac	643
Asp Arg Thr Ala Lys Leu Pro Arg Pro Gly Asp Thr Leu Leu Val Asp	
170 175 180	
cgc aaa gcg ggc tac gct ttt gag gcg att gcc aag acg gaa att tcg	691
Arg Lys Ala Gly Tyr Ala Phe Glu Ala Ile Ala Lys Thr Glu Ile Ser	
185 190 195	
agg ctt gcg ctg gaa gag gcg cca gat gtg tct tat cag gat att ggt	739
Arg Leu Ala Leu Glu Glu Ala Pro Asp Val Ser Tyr Gln Asp Ile Gly	
200 205 210	
ggc ttg gat gat cag att gaa ttg att caa gat gcc gtt gag ctg cca	787
Gly Leu Asp Asp Gln Ile Glu Leu Ile Gln Asp Ala Val Glu Leu Pro	
215 220 225	
ttt ttg cac ccg gag atg tac cgc gcc tac aac ctg cat cca cca aag	835
Phe Leu His Pro Glu Met Tyr Arg Ala Tyr Asn Leu His Pro Pro Lys	
230 235 240 245	
ggc gtg ctg ctg tac ggc cct ccc ggc tgt gga aag acg ctg att gct	883
Gly Val Leu Leu Tyr Gly Pro Pro Gly Cys Gly Lys Thr Leu Ile Ala	
250 255 260	
aag gct gtg gct aat tct ttg gcc aac cgc atc ggt gag act ggc acc	931
Lys Ala Val Ala Asn Ser Leu Ala Asn Arg Ile Gly Glu Thr Gly Thr	
265 270 275	
tcg tac ttc atc aac gtc aag ggg cca	958
Ser Tyr Phe Ile Asn Val Lys Gly Pro	
280 285	

&lt;210&gt; 78

&lt;211&gt; 286

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 78

```

Met Ser Ser Pro Thr Asp Ser Ser Pro Ser Asn Ser Phe Ser Asp Phe
 1           5           10           15

Asn Arg Glu Glu Gln Ser Arg Leu Ser Asp Glu Val Arg Gln Leu Lys
      20           25           30

Arg Thr Asn Ser Asp Leu Gly Ala Arg Asn Ala Lys Leu Ala Glu Met
      35           40           45

Leu Lys Ser Ser Arg Asp Lys Leu Ser Val Leu Phe Ser Gln Leu Glu
      50           55           60

Asp Met Ala Gln Pro Pro Ser Val Tyr Gly Thr Phe Leu Glu Thr Ala
      65           70           75           80

Lys Asp Gly Ser Asn Ala Glu Ile Phe Ala Gly Gly Arg Arg Met Arg
      85           90           95

Val Ala Val Ser Pro Met Leu Cys Ala Ala Asp Leu Met Pro Gly Val
      100          105          110

Gln Val Arg Leu Gly Glu Gly Asn Gln Val Leu Glu Ala Cys Asp Phe
      115          120          125

Glu Gln Thr Gly Glu Leu Ala Thr Leu Met Glu Met Ile Gly Arg Asp
      130          135          140

Arg Ala Leu Val Ser Asp Arg Ser Gly Glu Glu Arg Val Val Lys Leu
      145          150          155          160

Ala Gly Pro Leu Met Asp Arg Thr Ala Lys Leu Pro Arg Pro Gly Asp
      165          170          175

Thr Leu Leu Val Asp Arg Lys Ala Gly Tyr Ala Phe Glu Ala Ile Ala
      180          185          190

Lys Thr Glu Ile Ser Arg Leu Ala Leu Glu Glu Ala Pro Asp Val Ser
      195          200          205

Tyr Gln Asp Ile Gly Gly Leu Asp Asp Gln Ile Glu Leu Ile Gln Asp
      210          215          220

Ala Val Glu Leu Pro Phe Leu His Pro Glu Met Tyr Arg Ala Tyr Asn
      225          230          235          240

Leu His Pro Pro Lys Gly Val Leu Leu Tyr Gly Pro Pro Gly Cys Gly
      245          250          255

Lys Thr Leu Ile Ala Lys Ala Val Ala Asn Ser Leu Ala Asn Arg Ile
      260          265          270

Gly Glu Thr Gly Thr Ser Tyr Phe Ile Asn Val Lys Gly Pro
      275          280          285

```

&lt;210&gt; 79

&lt;211&gt; 735

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (101)..(712)

&lt;223&gt; RXA02317

&lt;400&gt; 79

```

gaattttccg gtgttaaaat gtgcggtgag cttggcggttg gcggaggacc agtgtgggga 60

gacgtcgaaa agcgaattca tggccccatc ttgccttaaa atg gcg cac atg cgc 115
                               Met Ala His Met Arg
                               1 5

tta ctg ctg acc tcc ttt ggc cat gat cat att cgg gat ttt gta cgc 163
Leu Leu Leu Thr Ser Phe Gly His Asp His Ile Arg Asp Phe Val Arg
                               10 15 20

ggt acc gtg gcg tat atc cct gat gcg acc agg ctt ttt gct gat agt 211
Gly Thr Val Ala Tyr Ile Pro Asp Ala Thr Arg Leu Phe Ala Asp Ser
                               25 30 35

ccc gag gct gct cct ttt atg gag acg gag cga aat atg ctg cgc gag 259
Pro Glu Ala Ala Pro Phe Met Glu Thr Glu Arg Asn Met Leu Arg Glu
                               40 45 50

cac ggc ttg agc att cgt gag ctg ccg att tcc acg tcg act ccg gag 307
His Gly Leu Ser Ile Arg Glu Leu Pro Ile Ser Thr Ser Thr Pro Glu
                               55 60 65

gaa gtg gat cgg gtg ctt ggt gag gtt gat ggg gtg tat gtg gcg ggc 355
Glu Val Asp Arg Val Leu Gly Glu Val Asp Gly Val Tyr Val Ala Gly
70 75 80 85

ggt gag act ttt gat ctg atg tgg ctg ctg cgt tcc aca ggc aat gat 403
Gly Glu Thr Phe Asp Leu Met Trp Leu Leu Arg Ser Thr Gly Asn Asp
90 95 100

gag gtg ttg att aag cat gtt cgc gct ggt cta ccg tat att gga acg 451
Glu Val Leu Ile Lys His Val Arg Ala Gly Leu Pro Tyr Ile Gly Thr
105 110 115

agc gcc ggc gcg gta att gca ggt cct tcg att gaa ccg atc agc ttt 499
Ser Ala Gly Ala Val Ile Ala Gly Pro Ser Ile Glu Pro Ile Ser Phe
120 125 130

ttg gat agc ccc gat gtc gcg ccg aat tta agc gac tat tca ggt cta 547
Leu Asp Ser Pro Asp Val Ala Pro Asn Leu Ser Asp Tyr Ser Gly Leu
135 140 145

ggc ctg tgc gag cat gtc gtg gtg ccc cat gct ggt ggc acg atc ccg 595
Gly Leu Cys Glu His Val Val Val Pro His Ala Gly Gly Thr Ile Pro
150 155 160 165

caa ttt ccc atc gat gtg ttt gcg gaa acc gtg cgc acc tac ggc gcc 643
Gln Phe Pro Ile Asp Val Phe Ala Glu Thr Val Arg Thr Tyr Gly Ala
170 175 180

gaa ttc ccg ctg gtc ctg ctt aaa gat gga cag gca ctg ctt atc gac 691
Glu Phe Pro Leu Val Leu Leu Lys Asp Gly Gln Ala Leu Leu Ile Asp
185 190 195

```

gac cac ggc gtc cac cta att taggatgggt ccccatgagc acc  
 Asp His Gly Val His Leu Ile  
 200

735

<210> 80  
 <211> 204  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 80  
 Met Ala His Met Arg Leu Leu Leu Thr Ser Phe Gly His Asp His Ile  
 1 5 10 15  
 Arg Asp Phe Val Arg Gly Thr Val Ala Tyr Ile Pro Asp Ala Thr Arg  
 20 25 30  
 Leu Phe Ala Asp Ser Pro Glu Ala Ala Pro Phe Met Glu Thr Glu Arg  
 35 40 45  
 Asn Met Leu Arg Glu His Gly Leu Ser Ile Arg Glu Leu Pro Ile Ser  
 50 55 60  
 Thr Ser Thr Pro Glu Glu Val Asp Arg Val Leu Gly Glu Val Asp Gly  
 65 70 75 80  
 Val Tyr Val Ala Gly Gly Glu Thr Phe Asp Leu Met Trp Leu Leu Arg  
 85 90 95  
 Ser Thr Gly Asn Asp Glu Val Leu Ile Lys His Val Arg Ala Gly Leu  
 100 105 110  
 Pro Tyr Ile Gly Thr Ser Ala Gly Ala Val Ile Ala Gly Pro Ser Ile  
 115 120 125  
 Glu Pro Ile Ser Phe Leu Asp Ser Pro Asp Val Ala Pro Asn Leu Ser  
 130 135 140  
 Asp Tyr Ser Gly Leu Gly Leu Cys Glu His Val Val Val Pro His Ala  
 145 150 155 160  
 Gly Gly Thr Ile Pro Gln Phe Pro Ile Asp Val Phe Ala Glu Thr Val  
 165 170 175  
 Arg Thr Tyr Gly Ala Glu Phe Pro Leu Val Leu Leu Lys Asp Gly Gln  
 180 185 190  
 Ala Leu Leu Ile Asp Asp His Gly Val His Leu Ile  
 195 200

<210> 81  
 <211> 774  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(751)  
 <223> RXA02644

cgtggggaag tggagccac actgaaaagc tttgcgaggt caacatggcc accaggctag 60

caa tac caa ctc cct cag gcc ggt caa gtc ttt gag cat gaa cta gag 163  
Gln Tyr Gln Leu Pro Gln Ala Gly Gln Val Phe Glu His Glu Leu Glu  
10 15 20

atc aag cgc tgc cga ttt ctg acc tat atc acg cgt gtg caa gat cag 211  
Ile Lys Arg Ser Arg Phe Leu Thr Tyr Ile Thr Arg Val Gln Asp Gln  
25 30 35

gag cag gct cgc gaa ttt att cac tcc atc aag gag ctg tat ccg gat 259  
Glu Gln Ala Arg Glu Phe Ile His Ser Ile Lys Glu Leu Tyr Pro Asp  
40 45 50

gcg cgt cat cat tgc agt gcc ttt att ttc cat gtg gat ggc tca aat 307  
Ala Arg His His Cys Ser Ala Phe Ile Phe His Val Asp Gly Ser Asn  
55 60 65

gat gtg gag cgt tcc tct gat gat ggc gaa cct tcc ggt acc gcg gga 355  
Asp Val Glu Arg Ser Ser Asp Asp Gly Glu Pro Ser Gly Thr Ala Gly  
70 75 80 85

aaa ccc atg cta gag gcg ttg cgt ggc tgc gga atg aaa gat att gcc 403  
Lys Pro Met Leu Glu Ala Leu Arg Gly Ser Gly Met Lys Asp Ile Ala  
90 95 100

gct gtg gtg gtg cgt tat ttc ggt ggc gta aaa ctg ggc act ggc gga 451  
Ala Val Val Val Arg Tyr Phe Gly Gly Val Lys Leu Gly Thr Gly Gly  
105 110 115

tta	gtt	aat	gcc	tac	acc	aac	gcg	gtg	acg	gag	ctt	cta	cct	gag	gtt	499
Leu	Val	Asn	Ala	Tyr	Thr	Asn	Ala	Val	Thr	Glu	Leu	Leu	Pro	Glu	Val	
		120					125					130				

ttg caa gtc acc cgc tct gtt cgg gag att ttc aag att gac ctt ccg 547  
Leu Gln Val Thr Arg Ser Val Arg Glu Ile Phe Lys Ile Asp Leu Pro  
135 140 145

cat tct gat gcg ggg cgc att gaa gcg aat ctg cgc ggc atg ggc atc 595  
His Ser Asp Ala Gly Arg Ile Glu Ala Asn Leu Arg Gly Met Gly Ile  
150 155 160 165

atc att act gac act gaa tat ggt gca gaa gtc aca tac acc ttg gct 643  
Ile Ile Thr Asp Thr Glu Tyr Gly Ala Glu Val Thr Tyr Thr Leu Ala  
170 175 180

tta ttg cct ggt gaa cag gct gcg gtg gaa tct caa ttg tca tcc atg 691  
Leu Leu Pro Gly Glu Gln Ala Ala Val Glu Ser Gln Leu Ser Ser Met  
185 190 195

atg ggt gca gaa att gaa ttg aaa gaa tcc ggg cac atg tgg gtg gaa 739  
Met Gly Ala Glu Ile Glu Leu Lys Glu Ser Gly His Met Trp Val Glu  
200 205 210

tcc ccg agt gac tagtgcggtg taagagcact aga 774  
Ser Pro Ser Asp

215

<210> 82  
 <211> 217  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 82  
 Met Ala His Leu Thr Gln Tyr Gln Leu Pro Gln Ala Gly Gln Val Phe  
   1                  5                  10                  15  
 Glu His Glu Leu Glu Ile Lys Arg Ser Arg Phe Leu Thr Tyr Ile Thr  
                   20                  25                  30  
 Arg Val Gln Asp Gln Glu Gln Ala Arg Glu Phe Ile His Ser Ile Lys  
           35                  40                  45  
 Glu Leu Tyr Pro Asp Ala Arg His His Cys Ser Ala Phe Ile Phe His  
   50                  55                  60  
 Val Asp Gly Ser Asn Asp Val Glu Arg Ser Ser Asp Asp Gly Glu Pro  
   65                  70                  75                  80  
 Ser Gly Thr Ala Gly Lys Pro Met Leu Glu Ala Leu Arg Gly Ser Gly  
                   85                  90                  95  
 Met Lys Asp Ile Ala Ala Val Val Val Arg Tyr Phe Gly Gly Val Lys  
           100                  105                  110  
 Leu Gly Thr Gly Gly Leu Val Asn Ala Tyr Thr Asn Ala Val Thr Glu  
   115                  120                  125  
 Leu Leu Pro Glu Val Leu Gln Val Thr Arg Ser Val Arg Glu Ile Phe  
   130                  135                  140  
 Lys Ile Asp Leu Pro His Ser Asp Ala Gly Arg Ile Glu Ala Asn Leu  
   145                  150                  155                  160  
 Arg Gly Met Gly Ile Ile Ile Thr Asp Thr Glu Tyr Gly Ala Glu Val  
           165                  170                  175  
 Thr Tyr Thr Leu Ala Leu Leu Pro Gly Glu Gln Ala Ala Val Glu Ser  
   180                  185                  190  
 Gln Leu Ser Ser Met Met Gly Ala Glu Ile Glu Leu Lys Glu Ser Gly  
   195                  200                  205  
 His Met Trp Val Glu Ser Pro Ser Asp  
   210                  215

<210> 83  
 <211> 1411  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(1411)  
 <223> RXN02820

&lt;400&gt; 83

```

caacgaaaca aatgcaagcc cccaatcatg ggtttctacc aattaaattt tctttcagaa 60
aatatctccc cacataaaaag ttccttgata ggctcgagag atg aaa gtg acc caa 115
                                         Met Lys Val Thr Gln
                                         1           5

agc aca ttc ctt aaa tcg gta gct gcg ttc act gtc gca gcc tta acc 163
Ser Thr Phe Leu Lys Ser Val Ala Ala Phe Thr Val Ala Ala Leu Thr
          10          15          20

ctg acc atc tct tcg tgt tcc agc ggt gaa gac acc tcc gca agc tcc 211
Leu Thr Ile Ser Ser Cys Ser Ser Gly Glu Asp Thr Ser Ala Ser Ser
          25          30          35

acg gat act gaa aac tcc tca acc caa gca gca gcg tct ccc cca ctt 259
Thr Asp Thr Glu Asn Ser Ser Thr Gln Ala Ala Ala Ser Pro Pro Leu
          40          45          50

gcg cct tgt gaa ctt ccc gcc gac gct tct gct gaa gag gaa gta gaa 307
Ala Pro Cys Glu Leu Pro Ala Asp Ala Ser Ala Glu Glu Glu Val Glu
          55          60          65

ggc act cac aca ggt gaa gat att tct gtt gcc ccg gaa atc ggt acc 355
Gly Thr His Thr Gly Glu Asp Ile Ser Val Ala Pro Glu Ile Gly Thr
          70          75          80          85

ggc tac cgc gag ggc atg acc cct gtt caa acc caa ggt tat gcg gtg 403
Gly Tyr Arg Glu Gly Met Thr Pro Val Gln Thr Gln Gly Tyr Ala Val
          90          95          100

gca act gca aac ccc atc gct tct gaa gca gcc tgc gcg gtg tta aga 451
Ala Thr Ala Asn Pro Ile Ala Ser Glu Ala Ala Cys Ala Val Leu Arg
          105          110          115

gaa ggc ggc act gca gct gat gct ctt gtc acc gcg cag ttt gtt ttg 499
Glu Gly Gly Thr Ala Ala Asp Ala Leu Val Thr Ala Gln Phe Val Leu
          120          125          130

gga ctg acg gaa ccg cag tcg tct ggc ctt ggt ggt ggc gga tac att 547
Gly Leu Thr Glu Pro Gln Ser Ser Gly Leu Gly Gly Gly Gly Tyr Ile
          135          140          145

ctg tac tac gac gcc gaa gcc aat gcg gtg aca gcc att gat ggc cgt 595
Leu Tyr Tyr Asp Ala Glu Ala Asn Ala Val Thr Ala Ile Asp Gly Arg
          150          155          160          165

gaa aca gcg cca gtt gct gct gat gaa aac tat ctc att cat gtt tct 643
Glu Thr Ala Pro Val Ala Ala Asp Glu Asn Tyr Leu Ile His Val Ser
          170          175          180

gca gag gat caa acg gca cct gtt cct gat gcc cga cgt tcc ggc agg 691
Ala Glu Asp Gln Thr Ala Pro Val Pro Asp Ala Arg Arg Ser Gly Arg
          185          190          195

tca att ggt gtg cca gga atc gtg gca gcc ctt gga cag ctg cat gat 739
Ser Ile Gly Val Pro Gly Ile Val Ala Ala Leu Gly Gln Leu His Asp
          200          205          210

tca ttc gga aag acc tcc tgg cag gac gtg ctg aca act ccg cag cag 787

```



Ser	Phe	Gly	Lys	Thr	Ser	Trp	Gln	Asp	Val	Leu	Thr	Thr	Pro	Gln	Gln		
215						220					225						
ctc	gca	act	gat	ggg	ttt	tcc	atc	agc	cct	cgc	atg	tca	gca	tca	att	835	
Leu	Ala	Thr	Asp	Gly	Phe	Ser	Ile	Ser	Pro	Arg	Met	Ser	Ala	Ser	Ile		
230					235					240					245		
gct	aac	tcc	gct	gag	gat	ctc	tcc	cac	gat	ccg	gaa	gct	gcc	gca	tat	883	
Ala	Asn	Ser	Ala	Glu	Asp	Leu	Ser	His	Asp	Pro	Glu	Ala	Ala	Ala	Tyr		
				250					255						260		
ttc	ctt	gat	gaa	aac	ggg	gat	gcg	aag	gca	ccc	ggc	aca	ctt	tta	caa	931	
Phe	Leu	Asp	Glu	Asn	Gly	Asp	Ala	Lys	Ala	Pro	Gly	Thr	Leu	Leu	Gln		
			265					270							275		
aac	cct	gac	tat	gca	gaa	acg	att	cgt	ctc	atc	tct	gaa	ggg	ggc	ccc	979	
Asn	Pro	Asp	Tyr	Ala	Glu	Thr	Ile	Arg	Leu	Ile	Ser	Glu	Gly	Gly	Pro		
		280						285							290		
gat	gcg	ttc	tac	acg	ggg	gag	att	gca	gca	gac	atc	gtg	gaa	cgc	gcc	1027	
Asp	Ala	Phe	Tyr	Thr	Gly	Glu	Ile	Ala	Ala	Asp	Ile	Val	Glu	Arg	Ala		
		295				300											
acc	cgt	gag	gtt	gac	ggg	ttc	aca	cca	tca	ctg	atg	agc	acg	gca	gat	1075	
Thr	Arg	Glu	Val	Asp	Gly	Phe	Thr	Pro	Ser	Leu	Met	Ser	Thr	Ala	Asp		
310					315					320					325		
ttg	gct	gcc	tac	act	ccg	gaa	act	cgt	gaa	gct	ttg	tgt	gct	ccc	tac	1123	
Leu	Ala	Ala	Tyr	Thr	Pro	Glu	Thr	Arg	Glu	Ala	Leu	Cys	Ala	Pro	Tyr		
				330					335						340		
cgc	gac	aag	att	gtt	tgt	ggc	atg	cca	ccg	tca	tca	tcg	ggg	ggc	gtc	1171	
Arg	Asp	Lys	Ile	Val	Cys	Gly	Met	Pro	Pro	Ser	Ser	Ser	Gly	Gly	Val		
		345						350						355			
aca	gtg	atg	gaa	acc	ctg	ggg	atc	ttg	aac	aac	ttt	gat	ctc	gcc	caa	1219	
Thr	Val	Met	Glu	Thr	Leu	Gly	Ile	Leu	Asn	Asn	Phe	Asp	Leu	Ala	Gln		
		360					365								370		
tac	cca	ccc	act	gag	gtt	ggg	ttg	gat	ggc	gga	ttg	cca	aat	gcg	gaa	1267	
Tyr	Pro	Pro	Thr	Glu	Val	Gly	Leu	Asp	Gly	Gly	Leu	Pro	Asn	Ala	Glu		
		375				380					385						
gct	gtt	cac	ctg	att	tca	gag	gct	gag	cgc	ctg	gct	tat	gct	gat	cgc	1315	
Ala	Val	His	Leu	Ile	Ser	Glu	Ala	Glu	Arg	Leu	Ala	Tyr	Ala	Asp	Arg		
390					395					400					405		
gat	gct	tac	atc	ggg	gat	cct	gct	ttc	gtg	gaa	gtt	cca	gca	ggg	ggg	1363	
Asp	Ala	Tyr	Ile	Gly	Asp	Pro	Ala	Phe	Val	Glu	Val	Pro	Ala	Gly	Gly		
				410					415						420		
gtc	caa	cag	tgg	atc	aac	cat	gtc	cac	acg	ggc	gaa	cac	tcc	aaa	ctt	1411	
Val	Gln	Gln	Trp	Ile	Asn	His	Val	His	Thr	Gly	Glu	His	Ser	Lys	Leu		
			425					430						435			

&lt;210&gt; 84

&lt;211&gt; 437

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 84

Met Lys Val Thr Gln Ser Thr Phe Leu Lys Ser Val Ala Ala Phe Thr  
 1 5 10 15  
 Val Ala Ala Leu Thr Leu Thr Ile Ser Ser Cys Ser Ser Gly Glu Asp  
 20 25 30  
 Thr Ser Ala Ser Ser Thr Asp Thr Glu Asn Ser Ser Thr Gln Ala Ala  
 35 40 45  
 Ala Ser Pro Pro Leu Ala Pro Cys Glu Leu Pro Ala Asp Ala Ser Ala  
 50 55 60  
 Glu Glu Glu Val Glu Gly Thr His Thr Gly Glu Asp Ile Ser Val Ala  
 65 70 75 80  
 Pro Glu Ile Gly Thr Gly Tyr Arg Glu Gly Met Thr Pro Val Gln Thr  
 85 90 95  
 Gln Gly Tyr Ala Val Ala Thr Ala Asn Pro Ile Ala Ser Glu Ala Ala  
 100 105 110  
 Cys Ala Val Leu Arg Glu Gly Gly Thr Ala Ala Asp Ala Leu Val Thr  
 115 120 125  
 Ala Gln Phe Val Leu Gly Leu Thr Glu Pro Gln Ser Ser Gly Leu Gly  
 130 135 140  
 Gly Gly Gly Tyr Ile Leu Tyr Tyr Asp Ala Glu Ala Asn Ala Val Thr  
 145 150 155 160  
 Ala Ile Asp Gly Arg Glu Thr Ala Pro Val Ala Ala Asp Glu Asn Tyr  
 165 170 175  
 Leu Ile His Val Ser Ala Glu Asp Gln Thr Ala Pro Val Pro Asp Ala  
 180 185 190  
 Arg Arg Ser Gly Arg Ser Ile Gly Val Pro Gly Ile Val Ala Ala Leu  
 195 200 205  
 Gly Gln Leu His Asp Ser Phe Gly Lys Thr Ser Trp Gln Asp Val Leu  
 210 215 220  
 Thr Thr Pro Gln Gln Leu Ala Thr Asp Gly Phe Ser Ile Ser Pro Arg  
 225 230 235 240  
 Met Ser Ala Ser Ile Ala Asn Ser Ala Glu Asp Leu Ser His Asp Pro  
 245 250 255  
 Glu Ala Ala Ala Tyr Phe Leu Asp Glu Asn Gly Asp Ala Lys Ala Pro  
 260 265 270  
 Gly Thr Leu Leu Gln Asn Pro Asp Tyr Ala Glu Thr Ile Arg Leu Ile  
 275 280 285  
 Ser Glu Gly Gly Pro Asp Ala Phe Tyr Thr Gly Glu Ile Ala Ala Asp  
 290 295 300  
 Ile Val Glu Arg Ala Thr Arg Glu Val Asp Gly Phe Thr Pro Ser Leu  
 305 310 315 320

Met Ser Thr Ala Asp Leu Ala Ala Tyr Thr Pro Glu Thr Arg Glu Ala  
 325 330 335

Leu Cys Ala Pro Tyr Arg Asp Lys Ile Val Cys Gly Met Pro Pro Ser  
 340 345 350

Ser Ser Gly Gly Val Thr Val Met Glu Thr Leu Gly Ile Leu Asn Asn  
 355 360 365

Phe Asp Leu Ala Gln Tyr Pro Pro Thr Glu Val Gly Leu Asp Gly Gly  
 370 375 380

Leu Pro Asn Ala Glu Ala Val His Leu Ile Ser Glu Ala Glu Arg Leu  
 385 390 395 400

Ala Tyr Ala Asp Arg Asp Ala Tyr Ile Gly Asp Pro Ala Phe Val Glu  
 405 410 415

Val Pro Ala Gly Gly Val Gln Gln Trp Ile Asn His Val His Thr Gly  
 420 425 430

Glu His Ser Lys Leu  
 435

&lt;210&gt; 85

&lt;211&gt; 507

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1) .. (507)

&lt;223&gt; FRXA02820

&lt;400&gt; 85

gct aac tcc gct gag gat ctc tcc cac gat ccg gaa gct gcc gca tat	48
Ala Asn Ser Ala Glu Asp Leu Ser His Asp Pro Glu Ala Ala Ala Tyr	
1 5 10 15	
ttc ctt gat gaa aac ggt gat gcg aag gca ccc ggc aca ctt tta caa	96
Phe Leu Asp Glu Asn Gly Asp Ala Lys Ala Pro Gly Thr Leu Leu Gln	
20 25 30	
aac cct gac tat gca gaa acg att cgt ctc atc tct gaa ggt ggc ccc	144
Asn Pro Asp Tyr Ala Glu Thr Ile Arg Leu Ile Ser Glu Gly Gly Pro	
35 40 45	
gat gcg ttc tac acg ggt gag att gca gca gac atc gtg gaa cgc gcc	192
Asp Ala Phe Tyr Thr Gly Glu Ile Ala Ala Asp Ile Val Glu Arg Ala	
50 55 60	
acc cgt gag gtt gac ggt ttc aca cca tca ctg atg agc acg gca gat	240
Thr Arg Glu Val Asp Gly Phe Thr Pro Ser Leu Met Ser Thr Ala Asp	
65 70 75 80	
ttg gct gcc tac act ccg gaa act cgt gaa gct ttg tgt gct ccc tac	288
Leu Ala Ala Tyr Thr Pro Glu Thr Arg Glu Ala Leu Cys Ala Pro Tyr	
85 90 95	
cgc gac aag att gtt tgt ggc atg cca ccg tca tca tcg ggt ggc gtc	336

Arg Asp Lys Ile Val Cys Gly Met Pro Pro Ser Ser Ser Gly Gly Val  
 100 105 110

aca gtg atg gaa acc ctg ggt atc ttg aac aac ttt gat ctc gcc caa 384  
 Thr Val Met Glu Thr Leu Gly Ile Leu Asn Asn Phe Asp Leu Ala Gln  
 115 120 125

tac cca ccc act gag gtt ggt ttg gat ggc gga ttg cca aat gcg gaa 432  
 Tyr Pro Pro Thr Glu Val Gly Leu Asp Gly Gly Leu Pro Asn Ala Glu  
 130 135 140

gct gtt cac ctg att tca gag gct gag cgc ctg gct tat gct gat cgc 480  
 Ala Val His Leu Ile Ser Glu Ala Glu Arg Leu Ala Tyr Ala Asp Arg  
 145 150 155 160

gat gct tac atc ggt gat cct gct ttc 507  
 Asp Ala Tyr Ile Gly Asp Pro Ala Phe  
 165

&lt;210&gt; 86

&lt;211&gt; 169

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 86

Ala Asn Ser Ala Glu Asp Leu Ser His Asp Pro Glu Ala Ala Ala Tyr  
 1 5 10 15

Phe Leu Asp Glu Asn Gly Asp Ala Lys Ala Pro Gly Thr Leu Leu Gln  
 20 25 30

Asn Pro Asp Tyr Ala Glu Thr Ile Arg Leu Ile Ser Glu Gly Gly Pro  
 35 40 45

Asp Ala Phe Tyr Thr Gly Glu Ile Ala Ala Asp Ile Val Glu Arg Ala  
 50 55 60

Thr Arg Glu Val Asp Gly Phe Thr Pro Ser Leu Met Ser Thr Ala Asp  
 65 70 75 80

Leu Ala Ala Tyr Thr Pro Glu Thr Arg Glu Ala Leu Cys Ala Pro Tyr  
 85 90 95

Arg Asp Lys Ile Val Cys Gly Met Pro Pro Ser Ser Ser Gly Gly Val  
 100 105 110

Thr Val Met Glu Thr Leu Gly Ile Leu Asn Asn Phe Asp Leu Ala Gln  
 115 120 125

Tyr Pro Pro Thr Glu Val Gly Leu Asp Gly Gly Leu Pro Asn Ala Glu  
 130 135 140

Ala Val His Leu Ile Ser Glu Ala Glu Arg Leu Ala Tyr Ala Asp Arg  
 145 150 155 160

Asp Ala Tyr Ile Gly Asp Pro Ala Phe  
 165

&lt;210&gt; 87

<211> 604  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (101)..(604)  
 <223> FRXA02000

<400> 87

```

caacgaaaca aatgcaagcc cccaatcatg ggtttctacc aattaaattt tctttcagaa 60
aatatctccc cacataaaaag ttccttgata ggctcgagag atg aaa gtg acc caa 115
                                     Met Lys Val Thr Gln
                                     1 5

agc aca ttc ctt aaa tcg gta gct gcg ttc act gtc gca gcc tta acc 163
Ser Thr Phe Leu Lys Ser Val Ala Ala Phe Thr Val Ala Ala Leu Thr
                10                15                20

ctg acc atc tct tcg tgt tcc agc ggt gaa gac acc tcc gca agc tcc 211
Leu Thr Ile Ser Ser Cys Ser Ser Gly Glu Asp Thr Ser Ala Ser Ser
                25                30                35

acg gat act gaa aac tcc tca acc caa gca gca gcg tct ccc cca ctt 259
Thr Asp Thr Glu Asn Ser Ser Thr Gln Ala Ala Ala Ser Pro Pro Leu
                40                45                50

gcg cct tgt gaa ctt ccc gcc gac gct tct gct gaa gag gaa gta gaa 307
Ala Pro Cys Glu Leu Pro Ala Asp Ala Ser Ala Glu Glu Glu Val Glu
                55                60                65

ggc act cac aca ggt gaa gat att tct gtt gcc ccg gaa atc ggt acc 355
Gly Thr His Thr Gly Glu Asp Ile Ser Val Ala Pro Glu Ile Gly Thr
                70                75                80                85

ggc tac cgc gag ggc atg acc cct gtt caa acc caa ggt tat gcg gtg 403
Gly Tyr Arg Glu Gly Met Thr Pro Val Gln Thr Gln Gly Tyr Ala Val
                90                95                100

gca act gca aac ccc atc gct tct gaa gca gcc tgc gcg gtg tta aga 451
Ala Thr Ala Asn Pro Ile Ala Ser Glu Ala Ala Cys Ala Val Leu Arg
                105                110                115

gaa ggc ggc act gca gct gat gct ctt gtc acc gcg cag ttt gtt ttg 499
Glu Gly Gly Thr Ala Ala Asp Ala Leu Val Thr Ala Gln Phe Val Leu
                120                125                130

gga ctg acg gaa ccg cag tcg tct ggc ctt ggt ggt ggc gga tac att 547
Gly Leu Thr Glu Pro Gln Ser Ser Gly Leu Gly Gly Gly Tyr Ile
                135                140                145

ctg tac tac gac gcc gaa gcc aat gcg gtg aca gcc att gat ggc cgt 595
Leu Tyr Tyr Asp Ala Glu Ala Asn Ala Val Thr Ala Ile Asp Gly Arg
                150                155                160                165

gaa aca gcg
Glu Thr Ala 604

```

<210> 88  
 <211> 168  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 88  
 Met Lys Val Thr Gln Ser Thr Phe Leu Lys Ser Val Ala Ala Phe Thr  
 1 5 10 15  
 Val Ala Ala Leu Thr Leu Thr Ile Ser Ser Cys Ser Ser Gly Glu Asp  
 20 25 30  
 Thr Ser Ala Ser Ser Thr Asp Thr Glu Asn Ser Ser Thr Gln Ala Ala  
 35 40 45  
 Ala Ser Pro Pro Leu Ala Pro Cys Glu Leu Pro Ala Asp Ala Ser Ala  
 50 55 60  
 Glu Glu Glu Val Glu Gly Thr His Thr Gly Glu Asp Ile Ser Val Ala  
 65 70 75 80  
 Pro Glu Ile Gly Thr Gly Tyr Arg Glu Gly Met Thr Pro Val Gln Thr  
 85 90 95  
 Gln Gly Tyr Ala Val Ala Thr Ala Asn Pro Ile Ala Ser Glu Ala Ala  
 100 105 110  
 Cys Ala Val Leu Arg Glu Gly Gly Thr Ala Ala Asp Ala Leu Val Thr  
 115 120 125  
 Ala Gln Phe Val Leu Gly Leu Thr Glu Pro Gln Ser Ser Gly Leu Gly  
 130 135 140  
 Gly Gly Gly Tyr Ile Leu Tyr Tyr Asp Ala Glu Ala Asn Ala Val Thr  
 145 150 155 160  
 Ala Ile Asp Gly Arg Glu Thr Ala  
 165

<210> 89  
 <211> 824  
 <212> DNA  
 <213> Corynebacterium glutamicum

<220>  
 <221> CDS  
 <222> (1)..(801)  
 <223> RXN03178

<400> 89  
 ccc acc acc gtg gtc acg ggc acg atg gaa gcc gcc aac atc gag ggc 48  
 Pro Thr Thr Val Val Thr Gly Thr Met Glu Ala Ala Asn Ile Glu Gly  
 1 5 10 15  
 tcc cgc gtg ggt gtc ggc gag gcc ggc cag tat acc gtt gat cag ctg 96  
 Ser Arg Val Gly Val Gly Glu Ala Gly Gln Tyr Thr Val Asp Gln Leu  
 20 25 30  
 ctg cac ggt ctt ctt tta gcc agc ggt aac gat gcg gcg tat atg ttg 144  
 Leu His Gly Leu Leu Leu Ala Ser Gly Asn Asp Ala Ala Tyr Met Leu

35	40	45	
gct cag gaa ctt ggt ggg gat caa gca acc ctg gag aaa gta aac gcg Ala Gln Glu Leu Gly Gly Asp Gln Ala Thr Leu Glu Lys Val Asn Ala 50 55 60			192
ctg gcc aag gag ttg ggc act caa gac acc ttc gtt gcc act tat tcc Leu Ala Lys Glu Leu Gly Thr Gln Asp Thr Phe Val Ala Thr Tyr Ser 65 70 75 80			240
ggt ttg gat gcg ccg gga atg tcg acc tcc gca tac gac atg tca ttg Gly Leu Asp Ala Pro Gly Met Ser Thr Ser Ala Tyr Asp Met Ser Leu 85 90 95			288
att tat cag cat gcg tgg cag aac ccg gtt ttc gag tcg att atc tcc Ile Tyr Gln His Ala Trp Gln Asn Pro Val Phe Glu Ser Ile Ile Ser 100 105 110			336
acc gat cac att gat ttc cct ggt tgg ggc gac aat gag ggt ttc caa Thr Asp His Ile Asp Phe Pro Gly Trp Gly Asp Asn Glu Gly Phe Gln 115 120 125			384
gtc tgg aac gat aac gcc ttg ttc atg aac gat cct gat ggc atc ggc Val Trp Asn Asp Asn Ala Leu Phe Met Asn Asp Pro Asp Gly Ile Gly 130 135 140			432
ggc aag acc ggc tac acc gac gac gcg aac cac acc ttt gtc ggc ggt Gly Lys Thr Gly Tyr Thr Asp Asp Ala Asn His Thr Phe Val Gly Gly 145 150 155 160			480
ctc gat cgg ggt ggt cgc cgc ctc gcc gcc gta ctc ttg gat tcc acc Leu Asp Arg Gly Gly Arg Arg Leu Ala Ala Val Leu Leu Asp Ser Thr 165 170 175			528
gtc agc gac att cgt ccg tgg gaa caa gca cga ttg ctt atc gac gcc Val Ser Asp Ile Arg Pro Trp Glu Gln Ala Arg Leu Leu Ile Asp Ala 180 185 190			576
tcc ctc ccc atc acg ccg ggg tcc ggc gtg ggc cag ctg ggc tcc ggc Ser Leu Pro Ile Thr Pro Gly Ser Gly Val Gly Gln Leu Gly Ser Gly 195 200 205			624
agc gcg aac gat gtg gca ccg gcg acc cca gaa tta cca gaa ccc acc Ser Ala Asn Asp Val Ala Pro Ala Thr Pro Glu Leu Pro Glu Pro Thr 210 215 220			672
gac aac ctg act tca ggt gag ggt ggg tcg cag aac acg ctt ctt aag Asp Asn Leu Thr Ser Gly Glu Gly Gly Ser Gln Asn Thr Leu Leu Lys 225 230 235 240			720
ctc gtg gtg ccc atc gga atc atc gtg ctg ttg cta atc gcc gca cta Leu Val Val Pro Ile Gly Ile Ile Val Leu Leu Leu Ile Ala Ala Leu 245 250 255			768
gcg tgg aca ttc aga tct ccc aag aaa aag aac taggtgttct tcttcacgac Ala Trp Thr Phe Arg Ser Pro Lys Lys Lys Asn 260 265			821
ctc			824

<210> 90  
 <211> 267  
 <212> PRT  
 <213> Corynebacterium glutamicum

<400> 90

```

Pro Thr Thr Val Val Thr Gly Thr Met Glu Ala Ala Asn Ile Glu Gly
  1              5              10              15

Ser Arg Val Gly Val Gly Glu Ala Gly Gln Tyr Thr Val Asp Gln Leu
              20              25              30

Leu His Gly Leu Leu Leu Ala Ser Gly Asn Asp Ala Ala Tyr Met Leu
              35              40              45

Ala Gln Glu Leu Gly Gly Asp Gln Ala Thr Leu Glu Lys Val Asn Ala
  50              55              60

Leu Ala Lys Glu Leu Gly Thr Gln Asp Thr Phe Val Ala Thr Tyr Ser
  65              70              75              80

Gly Leu Asp Ala Pro Gly Met Ser Thr Ser Ala Tyr Asp Met Ser Leu
              85              90              95

Ile Tyr Gln His Ala Trp Gln Asn Pro Val Phe Glu Ser Ile Ile Ser
              100              105              110

Thr Asp His Ile Asp Phe Pro Gly Trp Gly Asp Asn Glu Gly Phe Gln
              115              120              125

Val Trp Asn Asp Asn Ala Leu Phe Met Asn Asp Pro Asp Gly Ile Gly
              130              135              140

Gly Lys Thr Gly Tyr Thr Asp Asp Ala Asn His Thr Phe Val Gly Gly
  145              150              155              160

Leu Asp Arg Gly Gly Arg Arg Leu Ala Ala Val Leu Leu Asp Ser Thr
              165              170              175

Val Ser Asp Ile Arg Pro Trp Glu Gln Ala Arg Leu Leu Ile Asp Ala
              180              185              190

Ser Leu Pro Ile Thr Pro Gly Ser Gly Val Gly Gln Leu Gly Ser Gly
              195              200              205

Ser Ala Asn Asp Val Ala Pro Ala Thr Pro Glu Leu Pro Glu Pro Thr
              210              215              220

Asp Asn Leu Thr Ser Gly Glu Gly Gly Ser Gln Asn Thr Leu Leu Lys
  225              230              235              240

Leu Val Val Pro Ile Gly Ile Ile Val Leu Leu Leu Ile Ala Ala Leu
              245              250              255

Ala Trp Thr Phe Arg Ser Pro Lys Lys Lys Asn
              260              265

```

<210> 91  
 <211> 749  
 <212> DNA



&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(726)

&lt;223&gt; FRXA02859

&lt;400&gt; 91

cag tat acc gtt gat cag ctg ctg cac ggt ctt ctt tta gcc agc ggt	48
Gln Tyr Thr Val Asp Gln Leu Leu His Gly Leu Leu Leu Ala Ser Gly	
1 5 10 15	
aac gat gcg gcg tat ctg ttg gct cag gaa ctt ggt ggg gat caa gca	96
Asn Asp Ala Ala Tyr Leu Leu Ala Gln Glu Leu Gly Gly Asp Gln Ala	
20 25 30	
acc ctg gag aaa gta aac gcg ctg gcc aag gag ttg ggc act caa gac	144
Thr Leu Glu Lys Val Asn Ala Leu Ala Lys Glu Leu Gly Thr Gln Asp	
35 40 45	
acc ttc gtt gcc act tat tcc ggt ttg gat gcg ccg gga atg tcg acc	192
Thr Phe Val Ala Thr Tyr Ser Gly Leu Asp Ala Pro Gly Met Ser Thr	
50 55 60	
tcc gca tac gac atg tca ttg att tat cag cat gcg tgg cag aac ccg	240
Ser Ala Tyr Asp Met Ser Leu Ile Tyr Gln His Ala Trp Gln Asn Pro	
65 70 75 80	
gtt ttc gag tcg att atc tcc acc gat cac att gat ttc cct ggt tgg	288
Val Phe Glu Ser Ile Ile Ser Thr Asp His Ile Asp Phe Pro Gly Trp	
85 90 95	
ggc gac aat gag ggt ttc caa gtc tgg aac gat aac gcc ttg ttc atg	336
Gly Asp Asn Glu Gly Phe Gln Val Trp Asn Asp Asn Ala Leu Phe Met	
100 105 110	
aac gat cct gat ggc atc ggc ggc aag acc ggc tac acc gac gac gcg	384
Asn Asp Pro Asp Gly Ile Gly Gly Lys Thr Gly Tyr Thr Asp Asp Ala	
115 120 125	
aac cac acc ttt gtc ggc ggt ctc gat cgg ggt ggt cgc cgc ctc gcc	432
Asn His Thr Phe Val Gly Gly Leu Asp Arg Gly Gly Arg Arg Leu Ala	
130 135 140	
gcc gta ctc ttg gat tcc acc gtc agc gac att cgt ccg tgg gaa caa	480
Ala Val Leu Leu Asp Ser Thr Val Ser Asp Ile Arg Pro Trp Glu Gln	
145 150 155 160	
gca cga ttg ctt atc gac gcc tcc ctc ccc atc acg ccg ggg tcc ggc	528
Ala Arg Leu Leu Ile Asp Ala Ser Leu Pro Ile Thr Pro Gly Ser Gly	
165 170 175	
gtg ggc cag ctg ggc tcc ggc agc gcg aac gat gtg gca ccg gcg acc	576
Val Gly Gln Leu Gly Ser Gly Ser Ala Asn Asp Val Ala Pro Ala Thr	
180 185 190	
cca gaa tta cca gaa ccc acc gac aac ctg act tca ggt gag ggt ggg	624
Pro Glu Leu Pro Glu Pro Thr Asp Asn Leu Thr Ser Gly Glu Gly Gly	
195 200 205	
tcg cag aac acg ctg ctt aag ctc gtg gtg ccc atc gga atc atc gtg	672

Ser Gln Asn Thr Leu Leu Lys Leu Val Val Pro Ile Gly Ile Ile Val  
 210 215 220

ctg ttg cta atc gcc gca cta gcg tgg aca ttc aga tct ccc aag aaa 720  
 Leu Leu Leu Ile Ala Ala Leu Ala Trp Thr Phe Arg Ser Pro Lys Lys  
 225 230 235 240

aag aac taggtgttct tcttcacgac ctc 749  
 Lys Asn

<210> 92

<211> 242

<212> PRT

<213> Corynebacterium glutamicum

<400> 92

Gln Tyr Thr Val Asp Gln Leu Leu His Gly Leu Leu Leu Ala Ser Gly  
 1 5 10 15

Asn Asp Ala Ala Tyr Leu Leu Ala Gln Glu Leu Gly Gly Asp Gln Ala  
 20 25 30

Thr Leu Glu Lys Val Asn Ala Leu Ala Lys Glu Leu Gly Thr Gln Asp  
 35 40 45

Thr Phe Val Ala Thr Tyr Ser Gly Leu Asp Ala Pro Gly Met Ser Thr  
 50 55 60

Ser Ala Tyr Asp Met Ser Leu Ile Tyr Gln His Ala Trp Gln Asn Pro  
 65 70 75 80

Val Phe Glu Ser Ile Ile Ser Thr Asp His Ile Asp Phe Pro Gly Trp  
 85 90 95

Gly Asp Asn Glu Gly Phe Gln Val Trp Asn Asp Asn Ala Leu Phe Met  
 100 105 110

Asn Asp Pro Asp Gly Ile Gly Gly Lys Thr Gly Tyr Thr Asp Asp Ala  
 115 120 125

Asn His Thr Phe Val Gly Gly Leu Asp Arg Gly Gly Arg Arg Leu Ala  
 130 135 140

Ala Val Leu Leu Asp Ser Thr Val Ser Asp Ile Arg Pro Trp Glu Gln  
 145 150 155 160

Ala Arg Leu Leu Ile Asp Ala Ser Leu Pro Ile Thr Pro Gly Ser Gly  
 165 170 175

Val Gly Gln Leu Gly Ser Gly Ser Ala Asn Asp Val Ala Pro Ala Thr  
 180 185 190

Pro Glu Leu Pro Glu Pro Thr Asp Asn Leu Thr Ser Gly Glu Gly Gly  
 195 200 205

Ser Gln Asn Thr Leu Leu Lys Leu Val Val Pro Ile Gly Ile Ile Val  
 210 215 220

Leu Leu Leu Ile Ala Ala Leu Ala Trp Thr Phe Arg Ser Pro Lys Lys

225

230

235

240

Lys Asn

&lt;210&gt; 93

&lt;211&gt; 1212

&lt;212&gt; DNA

&lt;213&gt; Corynebacterium glutamicum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (101)..(1189)

&lt;223&gt; RXA00137

&lt;400&gt; 93

tttacagagt gcttatgagg caatcagcca ctaagtgttg agtaatctac tagtttggac 60

tagaagttac	ccactttcag	tgaattttta	aggagagaac	atg	gct	ttg	gct	gat	115
				Met	Ala	Leu	Ala	Asp	
				1				5	

acc	cga	ttt	gcc	act	cgt	cgt	cgc	gca	ctt	gcc	gca	aaa	ctg	gca	gct	163
Thr	Arg	Phe	Ala	Thr	Arg	Arg	Arg	Ala	Leu	Ala	Ala	Lys	Leu	Ala	Ala	
				10				15						20		

caa	cgg	atc	gac	tca	att	ttg	gtg	aca	agc	ccg	atc	cat	gtt	cgc	tat	211
Gln	Arg	Ile	Asp	Ser	Ile	Leu	Val	Thr	Ser	Pro	Ile	His	Val	Arg	Tyr	
			25					30					35			

ctc	agc	gga	ttc	acc	ggc	tcc	aac	ggc	gca	ctg	atc	gtg	aac	aaa	gat	259
Leu	Ser	Gly	Phe	Thr	Gly	Ser	Asn	Gly	Ala	Leu	Ile	Val	Asn	Lys	Asp	
		40					45					50				

ctc	tcc	gcg	cag	atc	tgc	acc	gac	ggc	cgc	tac	acc	acc	cag	atc	gca	307
Leu	Ser	Ala	Gln	Ile	Cys	Thr	Asp	Gly	Arg	Tyr	Thr	Thr	Gln	Ile	Ala	
		55					60					65				

gaa	gaa	gtc	ccg	gac	atc	gag	gcg	ctg	att	gag	cgt	gcc	tcg	gca	acg	355
Glu	Glu	Val	Pro	Asp	Ile	Glu	Ala	Leu	Ile	Glu	Arg	Ala	Ser	Ala	Thr	
	70				75					80					85	

acg	ctg	cta	gcg	cag	gtc	gaa	ggg	ccg	cgt	cgt	ata	gca	atc	gaa	gcc	403
Thr	Leu	Leu	Ala	Gln	Val	Glu	Gly	Pro	Arg	Arg	Ile	Ala	Ile	Glu	Ala	
				90					95					100		

gca	caa	acc	acc	ctg	gac	cag	cta	gac	agc	ctg	cgt	gaa	gca	acc	cag	451
Ala	Gln	Thr	Thr	Leu	Asp	Gln	Leu	Asp	Ser	Leu	Arg	Glu	Ala	Thr	Gln	
			105					110					115			

gaa	gac	gtc	gag	ctg	atc	ccc	gtg	tca	ggc	gtt	gtg	gaa	tcc	att	cgc	499
Glu	Asp	Val	Glu	Leu	Ile	Pro	Val	Ser	Gly	Val	Val	Glu	Ser	Ile	Arg	
		120					125						130			

ctg	acc	aaa	gac	agc	ttc	gaa	ctc	gac	cgc	ctc	cgc	gat	gtc	gca	gcg	547
Leu	Thr	Lys	Asp	Ser	Phe	Glu	Leu	Asp	Arg	Leu	Arg	Asp	Val	Ala	Ala	
		135				140					145					

ctg	gct	tcc	caa	gca	ttc	gaa	gat	tta	ctc	gca	gca	gga	gaa	ctc	gcc	595
Leu	Ala	Ser	Gln	Ala	Phe	Glu	Asp	Leu	Leu	Ala	Ala	Gly	Glu	Leu	Ala	

150	155	160	165	
gaa ggc cga tca gag cgc caa gtc gcc gcc gat ctg gaa tac cgc atg				643
Glu Gly Arg Ser Glu Arg Gln Val Ala Ala Asp Leu Glu Tyr Arg Met				
	170	175	180	
cgc ctg ttg gga gca gaa cgc ccc agc ttc gac acc atc gtg gcc tct				691
Arg Leu Leu Gly Ala Glu Arg Pro Ser Phe Asp Thr Ile Val Ala Ser				
	185	190	195	
gga cct aac tcc gcg aaa cca cac cac ggc gca ggc gac cgc atc ctc				739
Gly Pro Asn Ser Ala Lys Pro His His Gly Ala Gly Asp Arg Ile Leu				
	200	205	210	
cag cgc ggc gat cta gtc acc atc gat ttc ggc gca cac gca cgc gga				787
Gln Arg Gly Asp Leu Val Thr Ile Asp Phe Gly Ala His Ala Arg Gly				
	215	220	225	
ttc aac tcc gat atg acc cgc acc ctc gtt atg ggc gaa gca ggg gag				835
Phe Asn Ser Asp Met Thr Arg Thr Leu Val Met Gly Glu Ala Gly Glu				
	230	235	240	245
ttc gaa gca gaa atc tac gac atc gtc ctg cgc tcc caa ctc gct ggt				883
Phe Glu Ala Glu Ile Tyr Asp Ile Val Leu Arg Ser Gln Leu Ala Gly				
	250	255	260	
gtt gaa gca gcc tac tca ggc gcc aac ctc ttc gac atc gac gca gca				931
Val Glu Ala Ala Tyr Ser Gly Ala Asn Leu Phe Asp Ile Asp Ala Ala				
	265	270	275	
tgc cgc aaa atc atc gaa gac gca ggc tac ggc gaa tac ttc gtg cac				979
Cys Arg Lys Ile Ile Glu Asp Ala Gly Tyr Gly Glu Tyr Phe Val His				
	280	285	290	
tcc acc ggc cac ggc atc gga ctt gaa gtc cac gaa gcc cca agc gca				1027
Ser Thr Gly His Gly Ile Gly Leu Glu Val His Glu Ala Pro Ser Ala				
	295	300	305	
tcc aaa acc tca caa gga gtc cta gaa acc ggc tcc aca ctg acc atc				1075
Ser Lys Thr Ser Gln Gly Val Leu Glu Thr Gly Ser Thr Leu Thr Ile				
	310	315	320	325
gaa ccc gga att tac gtc ccc gga aag ggc ggc gta cgc atc gaa gac				1123
Glu Pro Gly Ile Tyr Val Pro Gly Lys Gly Gly Val Arg Ile Glu Asp				
	330	335	340	
acc ctg att att acc tca gga gca ccg gaa atc atc acc aag gtg agt				1171
Thr Leu Ile Ile Thr Ser Gly Ala Pro Glu Ile Ile Thr Lys Val Ser				
	345	350	355	
aag gac ctc atc gtg gtg taatctaggt gagctaateg gtc				1212
Lys Asp Leu Ile Val Val				
	360			

&lt;210&gt; 94

&lt;211&gt; 363

&lt;212&gt; PRT

&lt;213&gt; Corynebacterium glutamicum

&lt;400&gt; 94

Met Ala Leu Ala Asp Thr Arg Phe Ala Thr Arg Arg Arg Ala Leu Ala  
 1 5 10 15  
 Ala Lys Leu Ala Ala Gln Arg Ile Asp Ser Ile Leu Val Thr Ser Pro  
 20 25 30  
 Ile His Val Arg Tyr Leu Ser Gly Phe Thr Gly Ser Asn Gly Ala Leu  
 35 40 45  
 Ile Val Asn Lys Asp Leu Ser Ala Gln Ile Cys Thr Asp Gly Arg Tyr  
 50 55 60  
 Thr Thr Gln Ile Ala Glu Glu Val Pro Asp Ile Glu Ala Leu Ile Glu  
 65 70 75 80  
 Arg Ala Ser Ala Thr Thr Leu Leu Ala Gln Val Glu Gly Pro Arg Arg  
 85 90 95  
 Ile Ala Ile Glu Ala Ala Gln Thr Thr Leu Asp Gln Leu Asp Ser Leu  
 100 105 110  
 Arg Glu Ala Thr Gln Glu Asp Val Glu Leu Ile Pro Val Ser Gly Val  
 115 120 125  
 Val Glu Ser Ile Arg Leu Thr Lys Asp Ser Phe Glu Leu Asp Arg Leu  
 130 135 140  
 Arg Asp Val Ala Ala Leu Ala Ser Gln Ala Phe Glu Asp Leu Leu Ala  
 145 150 155 160  
 Ala Gly Glu Leu Ala Glu Gly Arg Ser Glu Arg Gln Val Ala Ala Asp  
 165 170 175  
 Leu Glu Tyr Arg Met Arg Leu Leu Gly Ala Glu Arg Pro Ser Phe Asp  
 180 185 190  
 Thr Ile Val Ala Ser Gly Pro Asn Ser Ala Lys Pro His His Gly Ala  
 195 200 205  
 Gly Asp Arg Ile Leu Gln Arg Gly Asp Leu Val Thr Ile Asp Phe Gly  
 210 215 220  
 Ala His Ala Arg Gly Phe Asn Ser Asp Met Thr Arg Thr Leu Val Met  
 225 230 235 240  
 Gly Glu Ala Gly Glu Phe Glu Ala Glu Ile Tyr Asp Ile Val Leu Arg  
 245 250 255  
 Ser Gln Leu Ala Gly Val Glu Ala Ala Tyr Ser Gly Ala Asn Leu Phe  
 260 265 270  
 Asp Ile Asp Ala Ala Cys Arg Lys Ile Ile Glu Asp Ala Gly Tyr Gly  
 275 280 285  
 Glu Tyr Phe Val His Ser Thr Gly His Gly Ile Gly Leu Glu Val His  
 290 295 300  
 Glu Ala Pro Ser Ala Ser Lys Thr Ser Gln Gly Val Leu Glu Thr Gly  
 305 310 315 320  
 Ser Thr Leu Thr Ile Glu Pro Gly Ile Tyr Val Pro Gly Lys Gly Gly

<400> 95														60	
tgccaacagg ggatatgccca ctgtgtaccc cacggcggatg atcgggtaaaaa tcctcgggcgc															
gcagatatattg ttotttgctgc tctaagggtga ttttttgggca														115	
											gtg	gtg	ggg	gtg	gtg
											Val	Val	Gly	Val	Val
											1				5
tcc acc cct gcg cgt aac ctg gga agc atg act aaa aca ctt ggt tcc														163	
Ser	Thr	Pro	Ala	Arg	Asn	Leu	Gly	Ser	Met	Thr	Lys	Thr	Leu	Gly	Ser
				10					15					20	
ctt cag ctg gaa gaa atc acg ctg acc ctc cct ctg act gaa gat gtg														211	
Leu	Gln	Leu	Glu	Glu	Ile	Thr	Leu	Thr	Leu	Pro	Leu	Thr	Glu	Asp	Val
			25						30				35		
gcc gat gaa cgc acc att gat gtg ttc gca cgc att gcc aca cgc gtc														259	
Ala	Asp	Glu	Arg	Thr	Ile	Asp	Val	Phe	Ala	Arg	Ile	Ala	Thr	Arg	Val
			40				45					50			
ggg ggg gaa gac ctt cca tat tta gta ttc ctg cag ggt ggg cct ggc														307	
Gly	Gly	Glu	Asp	Leu	Pro	Tyr	Leu	Val	Phe	Leu	Gln	Gly	Gly	Pro	Gly
	55					60					65				
aat gaa gct cca cgt cca agc ctt aat ccc ctc aac ccc aat tgg ttg														355	
Asn	Glu	Ala	Pro	Arg	Pro	Ser	Leu	Asn	Pro	Leu	Asn	Pro	Asn	Trp	Leu
						75				80					85
ggc gtg gcc ttg gag gaa tac cgc gtg gtc atg ttg gat caa cgt ggc														403	
Gly	Val	Ala	Leu	Glu	Glu	Tyr	Arg	Val	Val	Met	Leu	Asp	Gln	Arg	Gly
				90					95					100	
acc ggc cgt tcc acc cca gtg ggt aat gat att ttg gaa aaa ccc aca														451	
Thr	Gly	Arg	Ser	Thr	Pro	Val	Gly	Asn	Asp	Ile	Leu	Glu	Lys	Pro	Thr
			105					110					115		
gca gaa gta gtg gag tac tta tcc cac ctg cgc gca gat ggc att gtg														499	
Ala	Glu	Val	Val	Glu	Tyr	Leu	Ser	His	Leu	Arg	Ala	Asp	Gly	Ile	Val
			120				125					130			
cga gat gct gaa gcc ctg cgt aag cat ttg ggt gtg aat cag tgg aac														547	
Arg	Asp	Ala	Glu	Ala	Leu	Arg	Lys	His	Leu	Gly	Val	Asn	Gln	Trp	Asn
	135					140					145				

ctt tta ggc cag tcc ttc gga ggt ttc acc acc ctg cat tac ttg tcc	595
Leu Leu Gly Gln Ser Phe Gly Gly Phe Thr Thr Leu His Tyr Leu Ser	
150 155 160 165	
cgg cac gcc gat tcc ttg gac aac gtg ttt att acc ggc ggt ctc agc	643
Arg His Ala Asp Ser Leu Asp Asn Val Phe Ile Thr Gly Gly Leu Ser	
170 175 180	
gct att gat cgc cca gca gaa gac gtg tat gcc aac tgt tac aac cgc	691
Ala Ile Asp Arg Pro Ala Glu Asp Val Tyr Ala Asn Cys Tyr Asn Arg	
185 190 195	
atg cgc cga aac tct gag gaa ttc tac cgt cgc ttc ccg caa tta cgg	739
Met Arg Arg Asn Ser Glu Glu Phe Tyr Arg Arg Phe Pro Gln Leu Arg	
200 205 210	
gaa act ttc cga ggg ttg gtt aat cgt gct cgc gcc ggg gag att gtg	787
Glu Thr Phe Arg Gly Leu Val Asn Arg Ala Arg Ala Gly Glu Ile Val	
215 220 225	
ctt ccc acc ggc gaa gtt gtg tca gaa acc agg ctg cga tcc ctt ggt	835
Leu Pro Thr Gly Glu Val Val Ser Glu Thr Arg Leu Arg Ser Leu Gly	
230 235 240 245	
cac ttg ttg ggt agc aat gac ggc tgg ttt gat ctg tac aac ctg ctg	883
His Leu Leu Gly Ser Asn Asp Gly Trp Phe Asp Leu Tyr Asn Leu Leu	
250 255 260	
gaa tta gat ccc acc tcc aac gct ttt gtc cat gac ctg gca gga ctt	931
Glu Leu Asp Pro Thr Ser Asn Ala Phe Val His Asp Leu Ala Gly Leu	
265 270 275	
ttg cct ttc ggc aac cgc aac cca att tat tac gtg ctc cat gag tcc	979
Leu Pro Phe Gly Asn Arg Asn Pro Ile Tyr Tyr Val Leu His Glu Ser	
280 285 290	
tct tac gcc gac ggt gtg gtg aca aat tgg gca gca gag cgt gtg ctt	1027
Ser Tyr Ala Asp Gly Val Val Thr Asn Trp Ala Ala Glu Arg Val Leu	
295 300 305	
cca gag gat ttc cgc gag gat cca aca ctg ctc acc ggt gag cac gtg	1075
Pro Glu Asp Phe Arg Glu Asp Pro Thr Leu Leu Thr Gly Glu His Val	
310 315 320 325	
ttc cag gag tgg aca gac acc gtg ccg tcg ctc aag ccg tgg aag gac	1123
Phe Gln Glu Trp Thr Asp Thr Val Pro Ser Leu Lys Pro Trp Lys Asp	
330 335 340	
gtt gcc ctg gca ttg gct cag cag gaa tgg ccc aag ctt tat gat gcg	1171
Val Ala Leu Ala Leu Ala Gln Gln Glu Trp Pro Lys Leu Tyr Asp Ala	
345 350 355	
aag gca ttg gaa aac tca cag gcc aag ggc gct gca gca gtg tat gcc	1219
Lys Ala Leu Glu Asn Ser Gln Ala Lys Gly Ala Ala Val Tyr Ala	
360 365 370	
aat gac gtt ttc gtc cca gtg gat tac tct ctg gaa acc gca caa cac	1267
Asn Asp Val Phe Val Pro Val Asp Tyr Ser Leu Glu Thr Ala Gln His	
375 380 385	

ctg ccc ggt gtg cag ctg ttt atc acc agc cag cat gaa cac aat gga 1315  
 Leu Pro Gly Val Gln Leu Phe Ile Thr Ser Gln His Glu His Asn Gly  
 390 395 400 405

ctt cgt gcc agc tca ggc gca gta ctg aag cac ctt ttc gat ctg gcc 1363  
 Leu Arg Ala Ser Ser Gly Ala Val Leu Lys His Leu Phe Asp Leu Ala  
 410 415 420

cac ggc cga gag gta cgc tgattcctcg tgtagtact agc 1404  
 His Gly Arg Glu Val Arg  
 425

<210> 96

<211> 427

<212> PRT

<213> Corynebacterium glutamicum

<400> 96

Val Val Gly Val Val Ser Thr Pro Ala Arg Asn Leu Gly Ser Met Thr  
 1 5 10 15

Lys Thr Leu Gly Ser Leu Gln Leu Glu Glu Ile Thr Leu Thr Leu Pro  
 20 25 30

Leu Thr Glu Asp Val Ala Asp Glu Arg Thr Ile Asp Val Phe Ala Arg  
 35 40 45

Ile Ala Thr Arg Val Gly Gly Glu Asp Leu Pro Tyr Leu Val Phe Leu  
 50 55 60

Gln Gly Gly Pro Gly Asn Glu Ala Pro Arg Pro Ser Leu Asn Pro Leu  
 65 70 75 80

Asn Pro Asn Trp Leu Gly Val Ala Leu Glu Glu Tyr Arg Val Val Met  
 85 90 95

Leu Asp Gln Arg Gly Thr Gly Arg Ser Thr Pro Val Gly Asn Asp Ile  
 100 105 110

Leu Glu Lys Pro Thr Ala Glu Val Val Glu Tyr Leu Ser His Leu Arg  
 115 120 125

Ala Asp Gly Ile Val Arg Asp Ala Glu Ala Leu Arg Lys His Leu Gly  
 130 135 140

Val Asn Gln Trp Asn Leu Leu Gly Gln Ser Phe Gly Gly Phe Thr Thr  
 145 150 155 160

Leu His Tyr Leu Ser Arg His Ala Asp Ser Leu Asp Asn Val Phe Ile  
 165 170 175

Thr Gly Gly Leu Ser Ala Ile Asp Arg Pro Ala Glu Asp Val Tyr Ala  
 180 185 190

Asn Cys Tyr Asn Arg Met Arg Arg Asn Ser Glu Glu Phe Tyr Arg Arg  
 195 200 205

Phe Pro Gln Leu Arg Glu Thr Phe Arg Gly Leu Val Asn Arg Ala Arg  
 210 215 220